

**Spatio-temporal modelling for
issues in crime and security**

Toby Peter Davies

A thesis presented for the degree of
Doctor of Philosophy

University College London

Declaration

I, Toby Davies, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The distribution of incidents in time and space is a central issue in the study of crime, for both theoretical and practical reasons. It is also a context in which quantitative analysis and modelling has significant potential value: such research represents a means by which the implications of theory can be examined rigorously, and can also provide tools which support both policing and policy-making. The nature of the field, however, presents a number of challenges, particularly with regard to the incorporation of complex environmental factors and the modelling of individual-level behaviour. In this thesis, the techniques of complexity science are used to overcome these issues, and the approach is demonstrated using a number of examples from a range of crime types.

The thesis begins by presenting a network-based framework for the analysis of spatio-temporal clustering. It is demonstrated that signature ‘motifs’ can be identified in patterns of offending for burglary and maritime piracy, and that the technique provides a more nuanced characterisation of clustering than existing approaches. Analysis is then presented of the relationship between street network structure and the distribution of urban crime. It is shown that burglary risk is predicted by the graph-theoretic properties of street segments; in particular, those which correspond to levels of street usage. It is further demonstrated that the ‘near-repeat’ phenomenon in burglary displays a form of directionality, which can be reconciled with a novel street network metric. These results are then used to inform a mathematical model of burglary, which is situated on a network and which may be used for prediction. This model is analysed and its behaviour characterised in terms of urban form. Finally, a model is presented for a contrasting crime problem, the London riots of 2011, and used to examine a number of policy questions.

Acknowledgements

I wish to thank my supervisors, Professor Steven Bishop and Professor Shane Johnson, for their guidance throughout my doctoral research. Both have been consistently supportive and generous with the time and opportunities which they have offered to me, and I am extremely grateful for their advice. Thanks are also due to a number of others who have provided similar support in an informal capacity, in particular Professor Sir Alan Wilson and Professor Kate Bowers.

I also acknowledge the practical support of several agencies in the course of the research. West Midlands Police and the Metropolitan Police Service have both provided recorded crime data, along with more general assistance, and my collaboration with both forces has also been informative and helpful in a wider sense. I also acknowledge the financial support of the Engineering and Physical Sciences Research Council (EPSRC) via the Security Science Doctoral Training Centre, reference EP/G037264/1.

I am grateful for the hospitality afforded to me by the Centre for Advanced Spatial Analysis and Department of Security and Crime Science throughout the majority of my research. Both provided an enjoyable and productive working environment, and the support of colleagues in both institutions has been particularly valuable, most notably that of Hannah Fry, Peter Baudains, Adam Dennett, Helen Griffiths and Aiden Sidebottom.

In addition, I wish to thank a number of others whose support and friendship has been very meaningful during this period. Most prominent among these are Joe Hall, Matt Roberts, John Gibbons, Tim Grady, Daniel Mitchell, Daniel Hunter, Chloë Kirkham-Smith, Colette Owen-Rafferty and Emily Burton. Finally, any description of the contribution of my parents, Ann and Peter, would not do it justice, but theirs is foremost in my mind.

Contents

1	Introduction	15
1.1	Structure of the thesis	19
1.2	Background and motivation	21
1.3	Complexity science	26
1.3.1	Examples of complex systems	27
1.3.2	Techniques applied within complexity science	29
1.4	Complex networks	31
1.5	Summary	35
2	Characterisation of spatio-temporal clustering via network analysis	37
2.1	Introduction	37
2.2	Previous research examining crime clustering	42
2.3	Methods	44
2.3.1	Analytical setting	45
2.3.2	The Knox test	46
2.3.3	Event networks	48
2.3.4	Relationship with the Knox test	51
2.3.5	Network motifs	53
2.3.6	Motif interpretation	54
2.3.7	Event chains	56
2.3.8	Statistical analysis	59
2.3.9	Random network generation	63
2.4	Results	66
2.4.1	Data	67

2.4.2	Comparing datasets	67
2.4.3	Traditional clustering analysis	69
2.4.4	Identification of motifs	70
2.4.5	Analysis of chain length	78
2.5	Discussion	81
2.5.1	Crime patterns	82
2.5.2	Application in other domains	85
2.5.3	Extensions to the method	85

3 Quantifying the relationship between street network structure and burglary risk 87

3.1	Introduction	88
3.1.1	Crime and urban form	91
3.1.2	Shortcomings of existing approaches	98
3.1.3	Street network analysis	100
3.1.4	Mathematical representation	101
3.1.5	Physical analysis of street networks	105
3.1.6	Path-based metrics	106
3.1.7	Research using path-based metrics	111
3.2	Empirical analysis	112
3.2.1	Analytical strategy	112
3.2.2	Street network	117
3.2.3	Crime	126
3.2.4	General relationships	130
3.2.5	Regression analysis	133
3.3	Discussion	140
3.3.1	Relationships established	140
3.3.2	Practical implications	142
3.3.3	Implications for modelling	144
3.3.4	Further work	144

4	Network effects in repeat victimisation	147
4.1	Background	147
4.2	Theoretical background	150
4.3	Empirical analysis	156
4.3.1	Data and network representation	156
4.3.2	Clustering	158
4.3.3	Directionality	162
4.4	Discussion	185
4.4.1	The relationship between commonality and (near-)repeat of-fending	187
4.4.2	Implications for crime prevention and modelling	189
4.4.3	Wider applicability of commonality	191
5	A model for burglary on street networks	193
5.1	Introduction	193
5.1.1	Modelling rationale	194
5.1.2	Modelling approach	197
5.2	Review of previous modelling approaches	199
5.2.1	The Short model	207
5.2.2	Evaluation of the Short model	214
5.3	Proposed model description	217
5.3.1	Spatial setting	217
5.3.2	Modelling approach	220
5.3.3	Model dynamics	222
5.4	Model analysis	225
5.4.1	Initial numerical results	225
5.4.2	Algebraic analysis	229
5.4.3	Properties of the Laplacian	233
5.4.4	Implications for model stability	236
5.4.5	Relationship between static risk and equilibrium	238
5.4.6	Spectra of street networks	239

5.5	Model dynamics and community structure	243
5.5.1	Previous research	244
5.5.2	Analytical argument for community effects	247
5.5.3	Community effects in the burglary model	250
5.5.4	Numerical results	254
5.6	Discussion	261
5.6.1	Modelling outcomes	263
5.6.2	Mathematical outcomes	264
5.6.3	Real-world implications	266
5.6.4	Further work	267
6	A model of the London riots	269
6.1	Introduction	269
6.1.1	The London riots	272
6.1.2	Previous research on riots	274
6.2	Empirical analysis	276
6.2.1	Data	277
6.2.2	Riot characteristics	279
6.3	Model	285
6.3.1	Overall structure	287
6.3.2	Attractiveness	291
6.3.3	Riot participation	293
6.3.4	Spatial assignment	296
6.3.5	Interaction between police and rioters	298
6.3.6	Model integration	299
6.4	Results	300
6.4.1	Numerical simulations	301
6.4.2	Demonstration case	302
6.4.3	Susceptibility of sites	306
6.4.4	Police resources and response	308
6.5	Discussion	311

6.5.1	Findings	313
6.5.2	Future work	314
7	Summary and discussion	317
7.1	Summary	317
7.2	Unifying themes	321
7.2.1	Crime as a complex system	321
7.2.2	The importance of realistic environmental backcloth	323
7.2.3	Correspondence between theory and measurement	324
7.2.4	Alternative modelling approaches	325
7.2.5	Appropriate null models	327
7.3	Further work	328
7.3.1	Implementation and integration of predictive systems	328
7.3.2	Parameter selection	330
7.3.3	Effect of policing	331

Chapter 1

Introduction

The purpose of this thesis is to explore the application of techniques from the field of complexity science in the analysis and modelling of criminal phenomena. In particular, the thesis is concerned with the distribution of criminal events in time and space, and the way in which its understanding can be used to inform the prevention of crime. In the work which follows, it will be argued that traditional approaches are unable to account satisfactorily for a number of issues which play a crucial role in the study of crime: the physical structure of urban areas, for example, or the role of crowd effects in offending. The field of complexity science, however, provides a set of techniques which have been developed in order to address concerns of precisely this type, and the aim of this thesis is to demonstrate the utility of these approaches in the study of crime. This is done by examining a number of distinct topics, in which examples from several crime types are examined from a variety of perspectives. Each chapter builds on the work which precedes it, though the particular focus changes a number of times: the thesis begins with empirical pattern analysis, which is then refined in order to incorporate environmental factors, before the results of this analysis are used to inform predictive modelling. A number of significant methodological challenges arise within these topics, many of which concern the analysis of networks and their incorporation in models of social systems; the approaches used to overcome these constitute the primary technical contribution of the thesis.

The potential to inform and support crime prevention activity represents one of the primary motivations for undertaking criminological research. In general, insights concerning both how and why criminal events take place are of substantial practical value, since they provide a basis for the design of effective crime prevention strategies. Research concerning the spatio-temporal character of crime is a particularly clear example of this, since criminal events can be more readily disrupted if the times and places at which they occur are well-understood. Crucially, these aspects of crime are quantitative by nature, and therefore amenable to formal analysis and modelling of the type presented in this thesis. Such research can, however, take a number of forms. While the most basic objective of spatio-temporal research is simply to characterise the distribution of crime risk, it may also seek to establish the underlying mechanisms by which certain phenomena arise. This has the potential to inform more general principles of crime prevention and, extending the principle still further, can be used as the basis for predictive approaches. Examples of each of these types will be considered in the course of the thesis.

The thesis begins by considering the phenomenon of space-time clustering; that is, the tendency of criminal events to occur close to each other in both space and time. Clustering of this type is of particular significance within criminology, for both theoretical and practical reasons: not only can its existence be reconciled with theories of offender behaviour, but it also represents a natural focus for intervention. While the existence and extent of such clustering can be established using known statistical tests, such analysis offers little insight into the precise nature of the underlying patterns of events. In particular, significant qualitative differences can be observed between sets of events which are equivalent in terms of their macro-level clustering. In Chapter 2, it is demonstrated that the proximity of events can be represented by a series of networks, the structure of which can then be analysed in order to identify fine-scale ‘signatures’ of offending. In particular, ‘motif analysis’ is applied to these networks in order to identify common spatio-temporal configurations among small groups of clustered incidents. This approach, however, represents

a significant technical challenge: the nature of the networks considered means that the null models typically used in this type of analysis are not appropriate. This is overcome through the use of a novel network generation algorithm, and the subsequent analysis reveals significant targeting patterns in both residential burglary and maritime piracy.

The methods presented in Chapter 2 can be applied to any generic set of events occurring in time and space; that is, the analysis takes no account of the physical context in which the crimes in question are occurring. Many criminological theories, however, suggest that the physical environment plays a significant role in shaping offender behaviour, and Chapter 3 is concerned with the analysis of such effects. In particular, this chapter investigates the relationship between the street network and the risk of crime in urban areas, with particular focus on the crime of burglary. Since the street network is a primary determinant of the movement and awareness patterns of offenders, the existence of such a relationship is predicted by theory and, indeed, a number of studies have found evidence in support of this. The majority of these studies, however, examine aspects of network structure which do not correspond well to the real-world concepts invoked by relevant criminological hypotheses. It is shown in Chapter 3 that, by treating the entire street network as a complex entity, the application of sophisticated network metrics can provide a proper and objective measure of these concepts. In particular, it is demonstrated that network betweenness is a highly significant predictor of burglary risk, and the practical implications of this are discussed.

The relationship between the street network and crime is also the focus of Chapter 4; rather than considering the risk of crime over an extended period of time, however, the chapter examines the short-term phenomenon of (near-)repeat victimisation. This phenomenon is closely related to space-time clustering, and refers to the tendency - observed in many settings - for incidents to be followed closely by further offences in the near spatial vicinity. Although the spatial and temporal scales of

this effect are well-understood, one issue which is typically ignored in such analysis is that of directionality; that is, whether anything can be said about the locations of near-repeats, relative to the original incident, other than the fact that they are close. On the basis of theory, however, there is reason to expect that bias may be present in the locations of follow-up offences; specifically, that those street segments which share the highest proportion of traffic with the initially-victimised segment are most likely to be targeted. Since inter-segment relationships of this type are not captured by any existing network measure, a novel metric, referred to as commonality, is introduced for this purpose. By including this within a conditional logit model of target choice, it is shown that not only does follow-up victimisation display directionality, but also that commonality is a significant predictor of the risk to individual street segments. This has important practical implications, since it provides a means by which efforts to prevent follow-up offences can be more precisely directed. In addition, commonality itself has a number of interesting properties and is likely to be useful in the study of networks more generally.

The results of Chapters 3 and 4 represent clear evidence that the street network plays an important role in shaping the distribution of crime. Its omission from mathematical models of crime therefore represents a significant shortcoming, and Chapter 5 is concerned with proposing a new model - in this case, for burglary - which incorporates the street network explicitly. In the formulation presented, crime risk is modelled as a property of the links of a network, and the two fundamental concepts of risk heterogeneity and event dependence are encoded mathematically. The influence of the network is realised in two ways: the inherent risk on each segment is determined partly by its network centrality, and the diffusion of risk occurs only through the network. After demonstrating some of the properties of the model, it is then analysed formally, with particular focus on the relationship between the behaviour of the model and community structure in the underlying network. A number of analytic results are found, each of which has real-world interpretation in terms of, for example, hot-spot formation. Furthermore, community-level synchro-

nisation is also demonstrated in a stochastic implementation of the model, thereby extending the results of previous research concerning deterministic systems. The potential use of the model as the basis of a predictive system is then discussed.

The model presented in Chapter 5 concerns the crime of burglary, and the approach taken reflects the particular individual-level criminological concepts of relevance in that context. The range of crimes for which modelling is likely to be of value is wide, however, and Chapter 6 is concerned with its application to a crime which presents contrasting challenges: that of large-scale civil disorder, and the London riots of 2011 in particular. Disorder of this type has a number of features which distinguish it from burglary in theoretical terms: the role of micro-level environmental factors is less significant, for example, while offender behaviour is driven to a much greater extent by crowd effects. It therefore requires a different approach, demonstrates the wider power of modelling, and has application to policy. The chapter begins with basic empirical analysis of offender behaviour during the London disorder, via which a number of characteristics are identified. These insights are then used as the basis for the construction of a model, which describes the behaviour of offenders over the course of an episode of rioting. The model itself is a hybrid of existing models from other domains, such as epidemiology and retail modelling, which are put together by analogy with various stages of rioting. After it is calibrated in order to reproduce the general characteristics of the London disorder, the model is then used to address a number of real-world policy questions relating to the level and nature of police activity.

1.1 Structure of the thesis

The topics explored in the various chapters of the thesis are quite diverse, in terms of the particular aspects of crime considered and the techniques used, and the thesis has been structured with this in mind. In particular, material concerning previous literature and technical background has been placed at the location where it first becomes relevant.

Since the body of relevant literature varies considerably between topics, and is large overall, it will not be discussed in a stand-alone chapter. Each of the five substantive chapters will, instead, begin with a review of the previous research, and criminological theory, relevant to that particular topic. Because of the ordering of the chapters, this examination will be incremental, and much of the material covered will remain relevant beyond the chapter in which it is first introduced. A similar principle applies to Chapter 7, entitled Summary and Discussion: each substantive chapter will end with a discussion of the particular findings of that section of research, and Chapter 7 is reserved for discussion of unifying themes and a general overview of the material covered.

Since the majority of literature will be reviewed elsewhere, the remainder of this chapter will be used to provide a general background to the study of crime, and motivate the approach used in the remainder of this thesis. In the next section, the value of criminological research is discussed, together with the range of approaches which have typically been employed. Shortcomings of these approaches are identified, and it is argued that complexity science provides a number of tools via which these can be addressed. A brief review of relevant aspects of complexity science is then given, in which the notion of a complex system is outlined, and the techniques which are typically employed within the field are introduced.

One aspect of complexity science - the topic of network science - is discussed in significantly more detail towards the end of this chapter. Since networks feature so prominently throughout the thesis, and since basic concepts are frequently invoked, it is appropriate to introduce the relevant terminology and basic theory at the outset. All other technical material, however (including more sophisticated aspects of network theory), will be introduced as it becomes relevant.

With regard to notation, effort has been made to ensure that notation remains

consistent throughout the thesis, in particular when the same concept appears at several points. The number of terms used, however, means that there may be some cases in which the same symbol appears with varying meaning from chapter to chapter (though these should not be noticeable). It should be generally assumed that, although notation will always be consistent within chapters, there may be some variation between them.

1.2 Background and motivation

The study of crime is a field with an extensive history, and one which incorporates a diverse range of approaches and perspectives. In general, it is concerned with understanding how and why crime takes place, both as an academic exercise in its own right and as a means of establishing how it can be prevented. Historically, the field has been dominated by traditional criminological approaches, which concentrate on the notion of criminal disposition; in essence, such research seeks to determine what caused individuals to become involved in criminality. Among the many issues investigated in this context are such classical topics as the criminal justice system, individual psychology and the influence of societal conditions. This approach, however, is problematic in many respects, the most notable of which is its basic premise of criminality as an inherent character trait. Even those who commit crimes do not do so to the exclusion of all other activities, so the notion that criminality is an attribute that an individual either does or does not possess is fallacious.

Traditional approaches to the study of crime also suffer from a number of other, more practical, shortcomings. On one hand, the fact that many of the concepts involved are subjective, or defined vaguely, represents a barrier to rigorous empirical study. Perhaps more serious, though, is the difficulty associated with translating insights derived from this line of research into crime prevention policy. Even if individual-level causes of criminality could be identified, in many cases it is far from clear how they can be manipulated: effecting societal change, for example, is neither simple nor fast. It is in response to limitations such as these that more recent re-

search has examined crime from an alternative perspective. The central principle of the approach in question is to focus on the criminal event itself: the circumstances in which it took place, and the processes leading up to it. Regardless of the criminality of the offender, every crime involves such an event; crucially, though, to disrupt the event is substantially more feasible than to influence the offender. By applying measures which inhibit the commission of the crime itself - both by reducing opportunities and increasing the difficulty of those which do arise - the characteristics of the offender are rendered immaterial. This line of reasoning characterises the ‘situational’ approach to crime prevention (Clarke, 1980).

Although interventions of this nature can take many forms, one issue which is common to the majority of approaches is the need to understand where and when the events in question take place. By definition, for an event to be disrupted requires that it first be located, so that the effectiveness of situational measures depends upon them being targeted at places and times where crime has the potential to occur. This is one of the primary reasons for which a substantial volume of crime-related research is concerned with the spatio-temporal properties of crime: its distribution in space, its distribution in time, and the interaction between the two. As well as being of practical importance, such research is also notable within criminology for its explicitly quantitative nature: since the relevant aspects of crime are measured in ways that are both objective and quantitative, they represent natural candidates for analysis and modelling. Indeed, empirical research of this form has been particularly successful in identifying prevalent spatio-temporal phenomena that can be characterised quantitatively and that have immediate implications for prevention: repeat victimisation (Pease, 1998), hot-spot formation (Ratcliffe, 2004) and space-time clustering (Johnson *et al.*, 2007) are notable examples.

Common phenomena such as these also exemplify the relationship between spatio-temporal research and criminological theory. In many cases, including those listed above, the patterns observed can be interpreted as artefacts of principles thought to

govern the behaviour of offenders. Indeed, many criminological theories incorporate effects which ought to be reflected in the spatio-temporal distribution of crime, so that research of this type is crucial for both the validation of existing hypotheses and the formulation of new ones. Spatio-temporal issues are among the few aspects of theory which can be expressed in concrete and universal terms, and related research is therefore of particular value in this respect. Furthermore, the link with theory is also one of the reasons that the scope of spatio-temporal research is not limited to simply describing existing patterns: understanding *how* and *why* patterns arise, for example, may be of use in determining how they can be disrupted, and the possibility of forecasting crime locations is also of clear value. In many cases, issues such as these can be reconciled with criminological theories, the incorporation of which in formal quantitative research is one of the themes of this thesis.

One area of research which is worthy of particular note in this context is that of crime modelling: the building of abstract, simplified descriptions of criminal phenomena. In the sense used here, this refers specifically to models which are not purely statistical (*i.e.* based exclusively upon numerical association), but which involve, to some extent, the encoding of hypothesised processes in a mathematical (or, at least, quantitative) way. The value of such modelling lies in its ability to distil real-world concepts, which may be specified in vague terms, into well-defined mechanisms to which analytical tools can then be applied. As well as affording clarity, since it necessitates the expression of theoretical concepts in definite terms, the development of such models gives rise to a number of possibilities. In particular, they can be used to explore the effect of potential interventions in a quantitative way; that is, they provide a means by which the logical consequences of hypothetical manipulations can be expressed numerically. Because of this, such models are of great value as a means of informing policy, since they allow wide ranges of possible scenarios to be explored at low marginal cost (in comparison with real-world experiments, for example). This value is magnified still further if the model is of the type to which powerful analytical tools can be applied. In such scenarios, the fact that the model

is derived from theory is crucial, since this determines the extent of its applicability and allows refinements to be made in a well-motivated way.

Although the rationale for spatio-temporal analysis and modelling, as outlined above, is well-known, such work has historically received relatively little attention outside the criminological community. In particular, few attempts have been made to apply techniques from either mathematics or physics, and the majority of quantitative research has been statistical in nature. There are a number of reasons for this. Firstly, work of this type has traditionally suffered from a lack of credibility among police authorities and policy-makers, so that its potential real-world impact has been somewhat limited; this nullifies one of the key incentives to perform such research. Even purely academic work, though, has been discouraged by more practical difficulties. The difficulty of obtaining crime data in a form which is suitable for analysis, for example, has represented a significant impediment to research, and this problem also applies to contextual data. More significantly, though, there has also been a general opinion within the mathematical and physical sciences that it is simply not feasible to model crime in this way. The complexity of crime - the number of factors involved, and the fact that it concerns the actions of rational individuals - contrasts with the nature of most classical topics, and appears to prohibit its modelling by traditional approaches.

Recently, however, each of these concerns has been, to some extent, remedied. In ideological terms, the rise in prominence of the ‘crime science’ approach (see Laycock, 2005), which emphasises the need for research to be scientifically-grounded, has brought with it a greater recognition of the potential real-world value of modelling and other approaches. Simultaneously, the provision of data has also improved: all necessary information (including, crucially, geographical data) is now recorded systematically, in digital form, as standard. Similarly, the availability of contextual data, such as geographic and demographic information, is now greater than before. These factors, along with the significant modelling challenges presented, have

stimulated significant interest within the modelling community (see Gordon, 2010). Spatio-temporal models of a number of criminal phenomena have been built, including burglary (*e.g.* Short *et al.*, 2008), gang territoriality (*e.g.* Smith *et al.*, 2012) and ethnic violence (*e.g.* Lim *et al.*, 2007). These models vary considerably in focus: while some seek to explain pattern formation (*e.g.* Short *et al.*, 2008; Berestycki & Nadal, 2010), others examine the implications of criminological theory (*e.g.* Groff, 2007a; Birks *et al.*, 2012) or seek to replicate real-world scenarios in great detail (Malleson *et al.*, 2012). Although beyond the scope of this thesis, a substantial volume of research has also considered non-spatial issues: these include economic models (Gordon *et al.*, 2009), studies of optimal law enforcement (Garoupa, 1997), and work concerning criminal networks (Ballester *et al.*, 2006).

Many of the themes which arise in the work described above exemplify the challenges faced when modelling crime, and the techniques used to overcome them. To model crime comprehensively, for example, it is necessary to model individual-level behaviour in detail (Birks *et al.*, 2012), while also accounting for collective behaviour (Barthélemy *et al.*, 2010) and the influence of environment (Groff, 2007b). Although these aspects would traditionally have rendered such systems intractable, in terms of modelling, the development of tools designed specifically to be applied in such situations has allowed significant progress to be made. The field of ‘complexity science’ comprises an array of techniques which address a number of the main conceptual complications arising in real-world systems: rationality, for example, and the complex inter-dependence between entities. These apply to a broad class of systems from many fields, including many examples in economics and biology, and are particularly appropriate in the context of social systems (see Bellomo *et al.*, 2009). This thesis will be concerned primarily with the application of such techniques in the context of crime and, with this in mind, a brief review of the field will now be given.

1.3 Complexity science

Complexity science is concerned with addressing various features and phenomena which arise frequently in the study of large real-world systems and which complicate their analysis. The field can be considered from two perspectives: the array of techniques deployed within it, and the systems to which these are applied. The first of these is diverse and difficult to outline succinctly, primarily because the techniques in question have only one thing in common: their applicability to systems of a particular class. It is logical, therefore, to begin by defining this class of complex system before outlining the various approaches to its study. Unfortunately, this is not a straightforward task: the broad scope of the field comes at a cost in terms of the specificity with which it can be described. No universal definition of a complex system exists, and examples can be found in many disciplines.

As a result of this, the notion of a complex system is most often defined either by example or by specifying a number of general characteristics which a system must possess in order to be regarded as complex. Exactly which characteristics are stipulated depends on the context and focus of particular areas of research, but two properties are common to all working definitions (see Newman, 2011):

- 1) the system is composed of many interacting parts; and
- 2) the collective behaviour of those parts displays features which are not accounted for by the simple summation of their individual behaviours; *i.e.* the macro-level behaviour does not follow straightforwardly from the elementary mechanisms.

It is the second of these properties which represents the principal distinction between complex systems and others, and which mandates the use of purpose-built tools. Collective behaviours which arise in such a manner are referred to as ‘emergent’ phenomena, and this encompasses a number of non-linear effects. Many of these involve spontaneous ordering, and pattern formation in particular, but the definition also incorporates chaotic behaviour, which is itself a substantial topic of

study in the context of physical systems. The common attribute of all such phenomena is that they are ‘irreducible’, in the sense that they are only observed in the aggregate, and are not evident when the system is considered at a finer scale. This apparent conflict can arise for a number of reasons; regardless of its cause, however, its existence undermines classical approaches based on either simplification or mean-field approximation. Complexity science is fundamentally concerned with finding alternative means by which the two levels of behaviour can be reconciled.

As remarked above, one of the consequences of the breadth of scope of complexity science is that the discussion to this point has included a number of vague terms: it is unclear, for example, how many parts constitute ‘many’, and the definition of emergent behaviour is ambiguous. Indeed, precise definitions of these terms do not exist (or, at least, are not agreed upon) and their meaning is dependent upon context. In general, it has been argued that to identify complex systems is easier to define them, and that the nature of the field is best communicated by example (Johnson, 2007). With this in mind, the following section gives a brief (but certainly not exhaustive) outline of topics commonly studied within complexity science.

1.3.1 Examples of complex systems

Many of the attributes by which complex systems are defined are general characteristics, which are agnostic to the nature of the underlying system. As a result, examples can be found in a wide range of fields, and a number of those most relevant to the context of crime will be discussed.

Physical systems Many of the simplest examples of complex systems, and their distinctive behaviour, can be found in the study of physical systems. Examples of these include fluids, excitable media and physical computation systems, all of which involve fine-scale interactions. Indeed, one of the fundamental observations concerning complex systems was made by Anderson (1972), in the context of condensed matter physics. In that paper, it was argued that the notion of physical theories as universal laws was misguided, since systems obey different laws depend-

ing on the scale at which they are studied: equations for the state of a gas cannot be derived from quantum mechanics, for example. Such discrepancies are present in many simple systems, and several continue to be employed as canonical examples of emergent phenomena: the Ising (1925) model of electron spin is such an example.

Economies Although physical systems display many of the fundamental characteristics of complex systems, many commonly-studied systems also possess one additional feature: the presence of rational actors. This refers to any situation in which some of the components of the system act autonomously, and behave according to a decision process: one which maximises some notion of utility, for example. One such example, in which the concept of utility is particularly apt, is the study of economic systems. The field of ‘econophysics’ (see Mantegna & Stanley, 2000) is concerned with precisely this approach, and the relationship between micro- and macro-level economic phenomena. Financial markets are particular examples of systems which are large-scale, self-organised and open, and have been well-studied from this perspective (Johnson *et al.*, 2003), particularly in the context of catastrophic events (*e.g.* Haldane & May, 2011).

Ecology and biology The concept of rationality is also pertinent in the study of ecological systems. Such systems are typically composed of a large number of heterogeneous entities, and the long-term evolutionary processes involved provide a mechanism by which emergent phenomena can arise. Levin (1998) summarises the field from the perspective of complexity science, with particular reference to the ubiquity of feedback loops, self-organisation and path-dependence. Particular areas of study include systems of interacting species and their stability; indeed, a seminal paper on this topic by May (1972) anticipated many later developments in complexity science more generally. The issue of pattern formation is also one which arises regularly in the study of biological systems (see Winfree, 2001).

Human societies One further area of relevant research concerns human social systems: interactions between individuals, and their relationship with the physical environment. Interactions themselves, in the form of social networks, have been

studied in a number of contexts, and the recent availability of large digital datasets has facilitated research at an unprecedented scale. Such datasets have also been examined in the context of human mobility, and this has yielded a number of insights into the nature of urban activities. Indeed, the study of urban areas in general is a prominent research topic, and examines both their structure and function (Batty, 2007). This includes lower-scale study of topics such as segregation and traffic flow, but also the identification of scaling laws at the global level. Further work of note also includes that of Epstein (2006), which has focussed on the modelling of individual-level social phenomena.

Many of the properties displayed by complex systems can also be observed in systems of interest in the context of crime: many components (the individuals within a society) interact in a non-trivial way (autonomously, within a physical environment) and give rise to behaviour (crime patterns) which cannot be deduced simply from the micro-level behaviour. It is also the case that the latter two types of complex system - ecology and human societies - are of particular relevance: as will be demonstrated, ecological analogies inform several theories of offender behaviour, and the interaction of individuals in a complex environment will be a primary focus of this thesis.

1.3.2 Techniques applied within complexity science

Having introduced the notion of a complex system, the discussion of complexity science will now be completed by describing some of the techniques typically employed within it. Again, this will not be exhaustive, but will focus on those which are of greatest relevance to the study of human systems. Furthermore, one particular topic - the theory of networks - features particularly prominently within the thesis as a whole and will therefore be described in particular detail in a separate section.

Dynamical systems In general, describing the behaviour of a complex system involves two steps: modelling the behaviour of individual components (or groups of components), and coupling those models in a way which reflects the interactions

between components. When the quantities to be modelled evolve in continuous time, the first of these is typically achieved using differential equations, the properties of which are the main focus of the field of dynamical systems (Arrowsmith & Place, 1992). Given the focus on emergent phenomena, of particular interest is the study of non-linear phenomena, such as bifurcations and chaotic behaviour. These, and a number of similar concepts, are detailed extensively by Strogatz (1994).

Discrete dynamics An alternative approach to modelling individual component dynamics involves the consideration of time in discrete, rather than continuous, terms; that is, time is modelled as a series of ‘steps’ of fixed length. Although simple, this does correspond to some real-life situations, and models of this type provide some of the clearest examples of emergent behaviour, such as the chaotic behaviour of the logistic map (May, 1976). It is a small refinement to consider space in discrete terms also, and systems of this type - referred to as ‘cellular automata’ - are the subject of a substantial volume of research. The appeal of the approach lies in its transparency and the ability of very basic models to generate emergent phenomena (Berlekamp *et al.*, 2003), but it has also been shown to be applicable in a remarkable range of contexts. Indeed, it has been argued by Wolfram (2002) that the field may constitute a universal modelling paradigm.

Agent-based modelling Agent-based modelling is a particular form of discrete approach which relies on computer simulation as its primary tool. In such a model, the behaviour of individual entities is defined by rules, and the system is allowed to evolve from a given start point with no external manipulation. It constitutes a brute-force approach, with advantages and disadvantages. While models can be arbitrarily complex, and therefore correspond closely to real-world mechanisms, this comes at a cost in terms of the insight which can be gained from such models. Their value is perhaps most clear in testing the sufficiency of hypothesised mechanisms to generate desired outcomes, and Gilbert (2008) provides an introduction from the perspective of social science. The segregation model of Schelling (1971) is particularly well-known, and a number of further examples are presented by Epstein & Axtell (1996).

1.4 Complex networks

Complex networks are mathematical structures which encode relationships between sets of discrete entities. As such, they are particularly useful in the study of complex systems, since they provide a means by which the interactions between components of a system can be represented. Indeed, since the non-trivial configuration of these interactions may itself be the source of complexity, they play a central role in the study of complex systems.

The study of networks in the context of crime is one of the primary themes of this thesis, and a number of examples will be considered. In particular, both physical and abstract networks will be considered, representing the structure of urban streets and the proximity of crimes, respectively. Since the same basic theory and terminology is common to all cases, and will be invoked frequently, a brief introduction to core network concepts will now be given. This will cover the basic structures and notation used in all discussion of networks, while more specific concepts will be introduced throughout the thesis as they become relevant.

Network science is a substantial field of research in its own right, with a number of strands: empirical study of real-world networks, modelling of network formation, and applications in real-world models. Much of this research is not relevant in the present context, and a review will not be given; instead, relevant literature will be introduced at points throughout the thesis. A number of comprehensive reviews can be found elsewhere (Newman, 2003; Boccaletti *et al.*, 2006; Costa *et al.*, 2011), as can those which focus on the particular sub-fields of spatial (Barthélemy, 2011) and temporal (Holme & Saramäki, 2012) networks.

Mathematical representation

From a technical perspective, the term *network* refers to a mathematical object which is otherwise known as a *graph*. Indeed, the distinction between the two terms

is primarily one of context: ‘graph’ refers to the abstract object, whereas ‘network’ is more often used to imply that the structure in question represents real-world entities. The majority of terminology used to refer to networks is taken from the mathematical field of ‘graph theory’, which is an extensive and mature subject (see Bollobás, 2002).

Basic notation

A *network* $G = (V, E)$ is an ensemble of *vertices*, V , together with the *links*, E , which join them¹. The set of vertices, $V = \{v\}$, is a non-empty and countable set of N elements, and $E = \{e\}$ is a set of M elements, each of which is a pair of vertices. The number of vertices, N , is referred to as the *order* or the network, and the number of links as its *size*.

In order to be able to refer to elements of the network, its vertices are labelled using the integers $1, \dots, N$, where the order is unimportant as long as the labelling is consistent and unique. Each particular vertex is then referred to using a subscript, v_i , or the label itself, i , depending on which is convenient; the later convention will be adopted for the remainder of this discussion. In this way, specific links can be denoted using pair notation: if a link exists between two vertices, i and j , it is represented by the pair (i, j) . When a link is present, the vertices are said to be *adjacent* and each is a *neighbour* of the other.

At this point it is necessary to note a distinction between two types of network: *undirected* and *directed*. In an undirected graph, links have no orientation, and the ordering of vertices in any link (i, j) is unimportant. In the directed case, however, the ordering is significant, and such a link (i, j) is said to exist *from* i *to* j . Directed networks can contain links in both directions between a given pair of vertices, or just one. Graphically, directionality is usually represented by adding an arrow to a link; both undirected and directed examples are shown in Figure 1.1.

¹Vertices are also commonly referred to as *nodes*, and links as *edges*, but these terms are avoided here to avoid confusion with the concept of ‘edge effects’ and the use of ‘activity node’ within criminological theory.

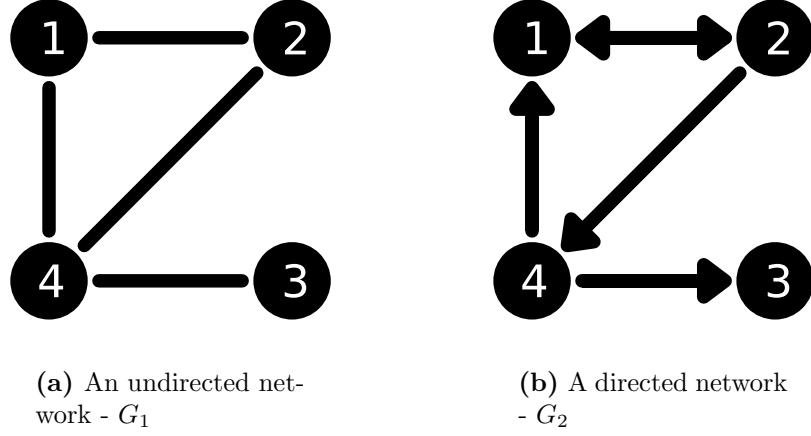


Figure 1.1: Graphical representations of simple undirected and directed networks. Circles represent vertices and the presence of a link is indicated by a line or arrow.

Two extreme cases, in terms of the presence of links, arise frequently and are identified by name. An *empty* network is one which contains no links (that is, $E = \emptyset$), and is simply defined by the number of vertices it contains. A *complete* network, on the other hand, is one in which all possible links are present. A complete undirected network therefore contains $\frac{N(N-1)}{2}$ links (the number of pairs of vertices), whereas a complete directed network contains $N(N-1)$ (since two links are present for each pair; one in each direction).

It is also possible for graphs (*i.e.* networks) to be derived from other networks. This can be done in a number of ways, many of which involve the notion of a *subgraph*. The vertices of a subgraph are a subset of the vertices of the original network, and its links are a subset of the links of those original network for which both vertices remain. Formally, then, for a network $G = (V, E)$, a subgraph $G' = (V', E')$ is one such that $V' \subseteq V$ and $E' \subseteq \{(i, j) \mid (i, j) \in E \text{ and } i, j \in V'\}$. An *induced subgraph* is a particular case in which all of the original links between members of V' are retained; that is, $E' = \{(i, j) \mid (i, j) \in E \text{ and } i, j \in V'\}$

Adjacency matrix

The structure of a network can be expressed in a number of ways, the most convenient and common of which is provided by the *adjacency matrix*. This encodes

all information necessary to describe a network: for the network, $G = (V, E)$, introduced above, the adjacency matrix, \mathbf{A} , is an $N \times N$ matrix such that

$$a_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E \text{ (i.e. there is a link connecting } i \text{ and } j) \\ 0 & \text{otherwise.} \end{cases} \quad (1.1)$$

The adjacency matrices of the two networks shown in Figure 1.1, for example, are

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad \mathbf{A}_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}. \quad (1.2)$$

For undirected networks, $a_{ij} = a_{ji}$ for all $i, j \in V$, and so these adjacency matrices are symmetric, with all information contained in the upper (or lower) triangular part. There is no such constraint for directed networks.

Network metrics

Many properties of network structure can be measured, and a number of metrics are common in the empirical study of networks. Each emphasises some different aspect of structure, and their use is dependent on the particular characteristics of interest. The number of metrics employed is very large, and only the most basic will be given here; a number of others, however, will be introduced in the course of the thesis.

Many of the quantities typically studied involve the properties of individual vertices, and many concern some notion of ‘centrality’. The most common concept is that of *degree*, which simply measures the number of links which are incident with a particular vertex (*i.e.* the number of neighbours it has). For a node i in an undirected network, this is typically denoted k_i . In terms of the adjacency matrix, \mathbf{A} , it is also straightforward to see that k_i is equal to $\sum_j a_{ij}$. In the directed case, analogous quantities of *in-degree* and *out-degree* are defined as the number of inward-pointing and outward-pointing links, respectively. The *degree distribution* is

the distribution of these values for a given network, and is frequently studied.

Other metrics concern larger-scale structures within networks, and the concept of *clustering* provides one such example. Clustering measures the tendency for links to form between the neighbours of a given node, and can be interpreted as the probability that the neighbours of a node should themselves be neighbours. For a given node, i , it is measured by the *clustering coefficient*, C_i , which is the ratio of the number of links between the neighbours of i to the maximum possible number of such links:

$$C_i = \frac{q_i}{k_i(k_i - 1)/2} \quad (1.3)$$

where q_i is the number of edges between neighbours of i , and k_i is the degree of i . Again, this can be averaged over all vertices of the network, providing a macro-level measure of clustering:

$$\langle C \rangle = \frac{1}{n} \sum_i C(i) \quad (1.4)$$

This is also an example of a metric which can, perhaps more intuitively, also be thought of in geometric or graphical terms. The clustering coefficient can also be expressed in terms of the closure of triangles in the network; that is, the probability that, when two sides of a triangle are already present, the third side will also be: It is notable that this quantity can be defined in several ways, typically involving the counting of certain types of subgraph:

$$C_{\Delta} = \frac{\text{number of closed triplets}}{\text{number of connected triples}}. \quad (1.5)$$

Graphical interpretations such as these will be used a number of times within the thesis.

1.5 Summary

The purpose of the argument presented in the latter part of this chapter was to motivate the study of criminal phenomena using the approach and tools of complexity science. The social systems in which crime occurs display many of the features

which were outlined as being characteristic of complex systems, and the modelling challenges involved are similar to those highlighted. Indeed, many of the contexts in which complex systems research has previously been carried out, such as ecology, are frequently invoked as analogies in the criminology literature. On this basis, the treatment of criminal phenomena in these terms is justified, and the following chapters will apply a number of the techniques identified in previous sections. The five substantial chapters, exploring a range of issues arising in the study of crime, will now follow.

Chapter 2

Characterisation of spatio-temporal clustering via network analysis

The primary subject of this chapter is the introduction of a novel method for the analysis of spatio-temporal event data, of the type commonly encountered during crime analysis. The work is motivated by the desire to characterise clustering phenomena in a more nuanced manner than previously possible, with the particular aim of developing the capability to identify *types* of clustering, rather than simply its existence. The terminology of complex networks provides a convenient and intuitive framework for clustering analysis, and allows existing approaches to be refined through the application of graph-theoretical techniques. Two main contributions are outlined: the introduction of the technique itself, along with a statistical approach specific to the particular context, and the application of the technique to real-world crime data, which reveals distinctive patterns in event distributions beyond those which have previously been observed empirically.

2.1 Introduction

The question of whether a set of events is clustered in space and time is a fundamental one in the analysis of crime, and one which is crucial to various aspects of theory and practice. As will be reviewed in more detail in Section 2.2, numerous empirical studies have shown that crime is distributed heterogeneously in space (*e.g.* Block *et al.*, 1995), and that patterns can also be observed in the times at which it

occurs, as typified by the existence of daily regularities and seasonal cycles (see, for example, Farrell & Pease, 1994). In both cases, such patterns can be reconciled with theoretical arguments concerning the behaviour of individuals and its relationship with crime (Cohen & Felson, 1979).

Such regularities are also of practical relevance, since areas or periods of disproportionate crime constitute an appealing target for interventions. ‘Hot-spot policing’ (Sherman *et al.*, 1989; Ratcliffe, 2004), whereby policing effort is concentrated on identified areas of high crime, is a straightforward example of this, and has been shown to be effective in reducing crime (Braga, 2001). More generally, the direction of police resources to those areas or situations where they have the greatest likelihood of interrupting or discouraging crime is likely to be an efficient strategy for deployment.

While clustering in either time or space is well-understood, though, a more subtle question concerns the interaction between the two dimensions; that is, whether there is dependence between the spatial and temporal separation of crimes. In its most immediate form, this corresponds to the notion that crimes which are close in space are more likely to be close in time (and *vice versa*), a relationship which is exemplified by the phenomenon of (near-)repeat victimisation (Pease, 1998; Bowers & Johnson, 2005). In situations where this is evident, the occurrence of an initial crime event implies an elevation of risk in the spatial vicinity for some period afterwards; that is, the spatial distribution of crime is dependent on recent events. Again, theoretical explanations have been offered for why this should be the case (Pease, 1998; Johnson & Bowers, 2004b), and the phenomenon has clear implications for crime prevention. In general, its existence implies that risk is communicable, in the sense that locations can be at increased risk of crime occurrence simply by virtue of their proximity to a recent victimisation.

This type of clustering can be tested for, and quantified, in several ways, the ma-

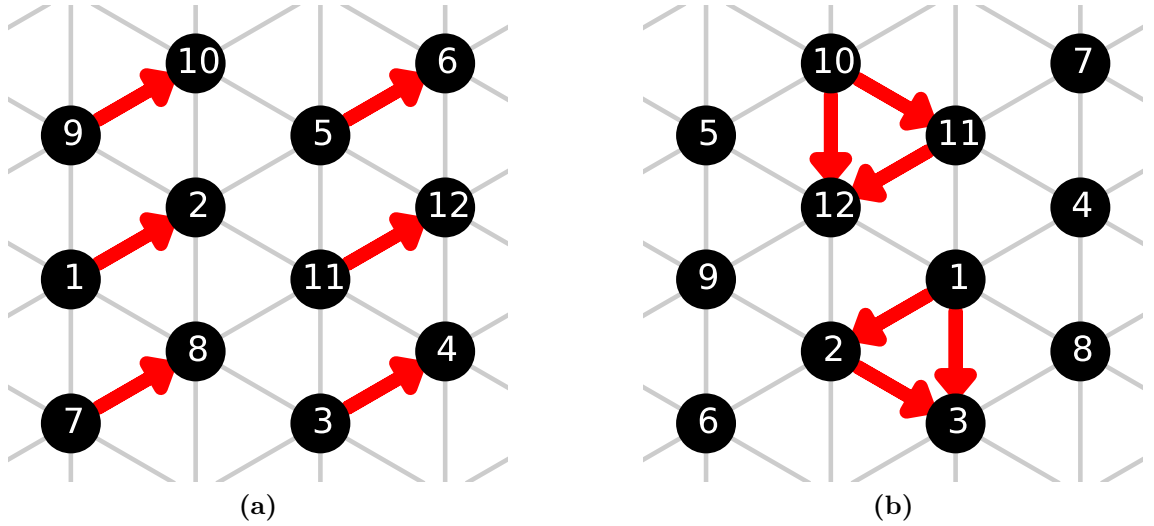


Figure 2.1: Close pair relationships for simple hypothetical sets of events. Vertex labels represent the time at which events occur, and their location is given by their position on the underlying grid, shown in grey. Red arrows represent close pairs, defined in this case as incidents occurring within two time units and one grid spacing.

jority of which involve comparison between the observed separation of events and that which would be expected if their spatial and temporal distributions were independent. Such methods typically involve the pair-wise comparison of events, where pairs are classified as being ‘close pairs’ if they lie within some specified thresholds in space and time (Knox, 1964; Mantel, 1967), or if they are nearest neighbours in space and time (Jacquez, 1996). These pair-wise relationships can then be used to characterise the data by, for example, comparing the number of close pairs against what would be expected if locations and timings were independent.

Although such techniques are certainly adequate to determine the existence of clustering *per se*, they are unable to provide any additional insight into the structure of the data. When defined on the basis of pair-counting only, the notion of clustering still allows for significant variability in the character of datasets, even amongst those which are found to be equivalently clustered. The hypothetical examples shown in Figure 2.1 illustrate the variation which is possible. Both datasets have the same number of close pairs, yet they exhibit perceptible qualitative differences: Figure 2.1a shows a series of isolated pairs, whereas two larger identifiable clusters are

present in Figure 2.1b. Since the marginal spatial and temporal distributions are identical in each case (the timings have simply been permuted), existing techniques would be unable to discriminate between the cases. The aim of the work presented here is to develop methods capable of identifying patterns such as these.

Considering Figure 2.1 in more depth, it can be seen that the essential difference between the cases is not in the number of close pairs, but rather their configuration. Any method capable of discerning between them should therefore consider the set of close pairs as a whole, and examine the relationships between them. A method by which this can be achieved arises from the observation that the dyadic ‘close pair’ relationship can be interpreted as defining a network. In this formulation, defined as an *event network*, events are represented as vertices and close pairs are joined by links, so that the analysis of the event data is translated to measurement of the network’s properties. Existing analytical techniques are trivially expressible in these terms, but the framework allows significantly more sophisticated techniques of network analysis, which have the potential to offer much greater insight into the structure of the underlying spatio-temporal data, to be applied.

One such technique with particular potential in this area is that of ‘motif analysis’ (Milo *et al.*, 2002). This refers to the identification of small subgraphs (connected groups of vertices) which occur with disproportionate frequency within a network: these are known as *motifs* since they represent recurring signatures in the composition of the network. Small structures such as these can be considered to be the fundamental ‘building blocks’ of the network, and thus reconciled with the micro-scale processes driving network formation. Analysis of this type has been successfully carried out in a variety of fields, including ecology and biology (*e.g.* Mangan & Alon, 2003), and also in a more abstract sense in the analysis of time series data (*e.g.* Xu *et al.*, 2008). The approach is particularly appropriate in this context since motifs have a clear interpretation in terms of the spatio-temporal patterns they represent, and can be interpreted as signatures of the targeting processes of criminal actors.

In addition, further analysis will consider the existence, and length, of *chains* within event networks. These represent sequences of events in which each successive pair of events is a close pair, and their length corresponds to the number of events which can be serially linked in this way. Features such as these have been considered in the context of criminal data in the past (*e.g.* Johnson & Braithwaite, 2009), and again have straightforward interpretation. Chain length can give an indication of the typical lifetime of an outbreak of events, or the characteristic ‘capacity’ of an area (the point at which continued offending is no longer profitable).

Measurement of both of these features - motifs and chains - presents a technical challenge due to the particular nature of event networks. Since the aim of the analysis is to identify patterns *over and above* the presence of clustering itself, it is necessary to control for such clustering when generating null distributions for comparison. In practical terms, this requires randomised networks to be matched in terms of the number of links present; however, while issues such as this are well-known in motif analysis, the methods typically used to resolve them cannot be applied in this case. Since their construction is geometric, event networks are constrained in the form they can take, and the link-rewiring methods typically employed to generate random networks for comparison are not guaranteed to produce valid event networks. Furthermore, in order to provide a meaningful comparison, the generation of such networks should correspond to random sampling from all possible sets of events. This is not the case for standard methods, for which the sampling is over all networks: the correspondence between event-sets and networks is not one-to-one. Because of these issues, bespoke methods are required to guarantee the validity of random network generation, and these will be introduced prior to the empirical work.

Application of the methods developed to real-world crime data demonstrates that they are indeed capable of identifying clustering phenomena at a higher resolution than previously possible. The method is applied to data for burglary and maritime

piracy, and spatio-temporal signatures can be clearly identified in both cases. These give rise to two key results: that the clustering present for these crimes is more dense than previously shown, and that the techniques developed are capable of discriminating between patterns which are indistinguishable by established approaches.

These results have considerable implications for work in crime analysis, which will ultimately be discussed. The emphasis of the chapter as a whole, however, is intended to be concentrated on the technique itself, rather than its specific application to crime. The reason for this is that, while crime analysis provides the motivation for its development, the method itself is agnostic to the nature of spatio-temporal data being considered, and could be applied in a number of other fields, such as epidemiology. With this versatility in mind, the theoretical context for these results in terms of criminal behaviour (*e.g.* routine activities, and ecological analogies) is not discussed in detail; indeed, since the precise context varies from crime to crime, the two examples considered here could not be discussed in a unified way. These criminological theories will be discussed fully in the particular context of burglary in the chapters which follow; here, however, the objective is simply to identify and characterise patterns of clustering.

2.2 Previous research examining crime clustering

Heterogeneity in the distribution of crime in space and time has been the subject of a substantial volume of research within criminology. Following notable early work by Shaw & McKay (1969), which established that offenders' residences tend to cluster in space, the fact that crime itself tends to be concentrated in certain locations has been established by a number of other studies. Sherman *et al.* (1989) found that a small number of areas account for a large proportion of predatory crime, and similar results were also found by Block & Block (1995) in the particular context of alcohol-related crime. Further studies have sought to relate these concentrations of crime to environmental and contextual factors (*e.g.* Cohen & Tita, 1999; Tita & Ridgeway, 2007).

The temporal dimension of crime has also received similar attention. Building on the work of Hawley (1950), routine activity theory (Cohen & Felson, 1979) suggests that the times at which crime occurs can be reconciled with the ‘rhythm’ and ‘tempo’ of life, at various scales. On one hand, crime concentration is known to vary by time of day (Ratcliffe, 2002; Felson & Poulsen, 2003; Thompson & Bowers, 2013), corresponding to changes in typical activity patterns. Over a longer scale, though, many crimes also exhibit seasonality (Farrell & Pease, 1994; Hipp *et al.*, 2004), and research has sought to reconcile this with the effect of environmental factors, such as temperature, on activity patterns (Field, 1992).

Although clustering in time and space is of considerable importance, more sophisticated recent work has focussed on the interaction between the two dimensions. The fact that the spatial distribution of crime varies over time can be seen at a number of scales: Lersch (2004) notes that some locations are prone to certain crime types at particular times of day, while the findings of Weisburd *et al.* (2004) demonstrate that the criminal character of places can vary over time. Similarly, Ratcliffe (2004) argues that crime ‘hot-spots’ can be classified according to a number of spatial and temporal criteria, such as their density and persistence in time. A number of methods have been proposed for the exploratory analysis of such patterns (*e.g.* Messner *et al.*, 1999; Ye & Wu, 2011).

Treating this issue more formally, a number of studies have sought to quantify the interaction between space and time statistically. Such interactions, of which space-time clustering is the most commonly-studied example, are indicative of phenomena such as repeat (Johnson *et al.*, 1997; Pease, 1998) or near-repeat (Morgan, 2001; Bowers & Johnson, 2005) victimisation and have significant practical implications. Studies which investigate this are typically performed using techniques originally proposed for epidemiology (*e.g.* Knox, 1964; Mantel, 1967; Jacquez, 1996) which test for inconsistency with a (null) hypothesis of independence of space and time.

These tests have been used to demonstrate the existence of clustering in a wide range of contexts, with burglary a particularly prominent example (Townnsley *et al.*, 2003; Johnson & Bowers, 2004a; Bowers & Johnson, 2005); indeed, work by Johnson *et al.* (2007) demonstrates that such patterns appear to be universal in data from several countries. Similar results have been found for other urban crimes, including shootings (Ratcliffe & Rengert, 2008), assault and robbery (Grubestic & Mack, 2008). Even more recently, the same techniques have been applied in less traditional contexts, with evidence for clustering also found in the locations of improvised explosive device attacks in Iraq (Johnson & Braithwaite, 2009) and incidents of maritime piracy (Marchione & Johnson, 2013).

Other techniques for the measurement of spatio-temporal interaction have also been proposed, notably the Markov chain method employed by Rey *et al.* (2012). Alongside this statistical work, other research has concentrated on investigating the stability of spatial concentrations of crime over time (Johnson *et al.*, 2008; Berk & Macdonald, 2009), including how such changes can best be visualised (Brunsdon *et al.*, 2007; Townnsley, 2008). Such research is distinct, though, from clustering analysis in the sense that is explored in this chapter, which is concerned explicitly with the dependence between individual events.

2.3 Methods

In this section, the methods by which network analysis can be used to characterise space-time clustering will be described. Before the network framework is introduced, though, one existing method - the Knox (1964) test - will be described. This provides a useful base for the work which follows in several respects: for example, it demonstrates how the consideration of ‘close pairs’ of incidents can be used to measure clustering, and highlights several of the methodological issues associated with this. Furthermore, it is the predominant method used in analysis of this type - particularly for crime - and, since the techniques which follow can be considered a

refinement of the general approach, it serves as a useful starting point.

After describing the Knox test, the network framework is introduced, and the relevance of motif- and chain-based analysis is explained. Finally, the challenges which arise in the analysis of such networks are identified, and the techniques by which they are overcome are described.

2.3.1 Analytical setting

Since a number of concepts, such as the form of data and measurements required, are common to all the methods which will subsequently be described, it is logical to outline the fundamental question which is to be addressed, and to fix the notation which will be used throughout.

In general, all methods apply to any generic set of events occurring in space and time. It is assumed that there are N such events, indexed by i , each of which has some spatial location x_i and a time of occurrence t_i . The spatial location can take any form (*e.g.* latitude/longitude or easting/northing), provided that there is an associated distance metric, d , for the separation between any pair of events. The shorthand d_{ij} will be used to represent the distance between two events i and j , so that

$$d_{ij} = d(x_i, x_j) \geq 0 \tag{2.1}$$

Similarly, $t_{ij} = t_j - t_i$ is defined as the temporal separation between events, though in this case the value can be positive or negative, depending on whether j occurs after or before i , respectively. Although only the absolute value will be used in some contexts (*e.g.* the Knox test), the distinction is significant in others.

Only inter-event distances are relevant throughout the analyses, so that the d_{ij} and t_{ij} values encode all relevant information for any given set of events. All techniques considered here seek to address the question of how to characterise the clustering patterns present in sets of events of this form.

2.3.2 The Knox test

The Knox (1964) test has been used widely in studies which address the simple question of whether there is interaction between the spatial and temporal distributions of a set of incidents; *i.e.* whether space-time clustering is present. For simplicity, and in keeping with the work which follows, a basic formulation will be described here, although more complex variations have been used elsewhere (and, indeed, will be considered in Chapter 4).

The origin of the test lies in epidemiology; more specifically, in the study of childhood leukaemia. Its motivation is to test whether a set of event data (in this context, observed cases of disease) is consistent with having been generated by a process involving an element of contagion, as is a common hypothesis in many such scenarios. The underlying rationale is that, if contagion is present, events will tend to be followed by other events in some spatial vicinity more than would be expected on the basis of chance. Such a relationship corresponds to a dependence between the spatial and temporal distributions: events are more likely to be close in space when they are close in time, and *vice versa*.

The basis for the test is the concept of a ‘close pair’ of events: one for which the the spatial and temporal separations of the events both lie within certain thresholds. For concreteness, D will be taken to represent the spatial threshold and T the critical temporal separation. These thresholds can be taken to have any value, but are typically selected on the basis of the anticipated radius of any contagion effect (which, in turn, may be informed by other analysis). More sophisticated versions of the test consider several bands in each dimension, which represent different levels of separation (see Chapter 4); for simplicity, though, only the simple single-threshold case is considered here.

The first step of the test is to compare every possible pair of events in the dataset (for N events, there will be $\frac{N(N-1)}{2}$ comparisons) and to record the number of those which

are close pairs (*i.e.* the number of pairs $\{i, j\}$ for which $d_{ij} \leq D$ and $-T \leq t_{ij} \leq T$). This statistic, denoted S_K , represents the observed proximity of events, as defined by the binary close pair relationship.

Once this count of close pairs has been found, it must be compared against what would be expected under the null hypothesis (that the events' locations in time and space are independent). In Knox's original work, it was assumed that the number of close pairs followed a Poisson distribution, and that the expected frequency could be computed using the marginal frequencies of spatial close pairs and temporal close pairs.

An alternative method, however, which has been used in recent work concerned with crime (*e.g.* Johnson *et al.*, 2007), is to employ a Monte Carlo approach. Rather than computing a theoretical value, this involves examining the values of the statistic in question (in this case, the number of close pairs) for a number of explicitly-constructed alternative datasets, generated under the assumption of the null hypothesis.

There are a number of ways in which the construction of these datasets can be performed, but a popular one is to use a permutation approach. Starting with the observed event data, sets of randomised events are generated by repeatedly permuting the timings of events, while maintaining the spatial information. Denoting the permutation at a given iteration by σ , the pair-wise comparison between two events i and j therefore involves comparison of their true spatial locations x_i and x_j (as before) but their permuted time-points $t_{\sigma(i)}$ and $t_{\sigma(j)}$. In other words, the temporal components of the events are shuffled, so that any alignment with the spatial components will be broken down. These sets of events therefore correspond to what would be expected under the null hypothesis: if there is no association between spatial and temporal distributions, the shuffling ought to make no significant difference to the number of close pairs observed.

The permutation approach is also appealing in another respect, which is that it is based entirely on observed data. Rather than synthesising events, the observed information is simply restructured; for each event set constructed, therefore, the spatial and temporal distributions are identical to the observed case. Anomalous results cannot, therefore, be ascribed to a change in either of the marginal distributions.

Given the method of generating randomised sets of events - the permutation of one dimension - the remaining analysis is simple. A number, n_K , of sets of events are constructed in this way, and the statistic of interest (the number of close pairs) is computed in each case; this is denoted \tilde{S}_K for the shuffled data. This can then be used as a reference distribution against which the true observed value can be compared.

The deviation from this distribution can be quantified: if r_K is the rank at which S_K would appear in an ordered list of the \tilde{S}_K values generated, then, as proposed by North *et al.* (2002), a pseudo-significance is given by:

$$p = \frac{r_K}{n_K + 1}. \quad (2.2)$$

Furthermore, the magnitude of the effect can be estimated by computing the z -score of S_K , relative to the null distribution. Significant deviation from the reference distribution indicates that it is improbable that the observed data could have been generated if there was no association between the timings and locations of events.

2.3.3 Event networks

The Knox test demonstrates how the concept of ‘close pairs’, and the analysis of their frequency, can be used to characterise space-time clustering. This inspires the fundamental innovation of this work, which is to use that relationship to define a network on the events considered. Such networks are, in essence, simply a

re-encoding of the set of close pair relationships, but the expression as a network provides a convenient instrument for various forms of clustering analysis.

As with the Knox test, the analysis assumes the existence of a spatial radius, D , and temporal radius, T , which are taken to define what it is for two events to be ‘close’ in either dimension. These are, of course, arbitrary, and therefore act as parameters for the construction of proximity networks; their variation is explored in the empirical work in the later part of the chapter. For the purpose of describing the method, though, it is assumed that these values are fixed.

The close pair relationships can be used to define two networks: one, G_d^D , on the basis of spatial proximity, and another, G_t^T , on the basis of temporal proximity. In both cases, the vertices represent the events in the dataset; vertex i corresponds to the i th event.

G_d^D is an undirected network, where any two events are connected by a link if they occurred within a distance D of each other. Formally, then, its set of links is given by

$$E_d^D = \{(i, j) \mid d_{ij} \leq D\}. \quad (2.3)$$

Since there is temporal ordering in the data, however, G_t^T is taken to be a directed network. In this case, if two events occurred within a time window T , a link is added from the earlier event to the later, so that the link-set, E_t^T , is:

$$E_t^T = \{(i, j) \mid 0 < t_{ij} \leq T\}. \quad (2.4)$$

It is assumed, in general, that the temporal measurements are sufficiently granular that, for any pair of events, one can always be said to have occurred first. Where this is not possible, however, cases for which $t_{ij} = t_{ji} = 0$ are resolved by deeming one event to have occurred first, at random, and directing the links in accordance with this.

It is worth noting, at this stage, that the graph G_d^D is an example of a particular type of *geometric graph* (Pach, 2004) known as a *unit disk graph* (Clark *et al.*, 1990). Such a graph is one whose links correspond to the intersections between circles in a plane; in this case, the circles of radius D around each event location. Since graphs constructed in this way are subject to geometric constraints, they have particular properties and require bespoke analysis, examples of which are presented by Marathe *et al.* (1995). A related field, relevant to the analysis which follows, is that of *random geometric graphs*, in which graphs are derived from the proximity of points distributed randomly in space, and a number of theoretical results have been derived for these (Penrose, 2003). Although many of these are not of immediate relevance, the fundamental issue of the analysis is similar: that the randomness lies in the positioning of the points, rather than explicitly in the network itself.

From the perspective of close pairs, G_d^D and G_t^T carry all relevant information about the set of events. Trivially, pairs of events which are close in both space *and* time can be identified by consulting the two networks: events i and j are close in space and time if the corresponding vertices are adjacent in both G_d^D and G_t^T . This can be formulated explicitly by defining the network G_{dt}^{DT} , which is the directed network of pairs which are close in both space and time. This is the *event network* for the dataset, and summarises the proximity relationships with which the analysis of space-time clustering is concerned. It is the intersection of the spatial and temporal networks defined previously, so that

$$E_{dt}^{DT} = \{(i, j) \mid (i, j) \in E_d^D \quad \textbf{and} \quad (i, j) \in E_t^T\}. \quad (2.5)$$

The relationship between the three networks G_d^D , G_t^T and G_{dt}^{DT} is shown visually in Figure 2.2a. The overall event network G_{dt}^{DT} is derived from the other two; however, it is important to note that it does not include the full extent of known information about the set of events. If a pair of events is not linked in G_{dt}^{DT} , for example, it is impossible to know whether that is because the link is absent in G_d^D , in G_t^T , or both.

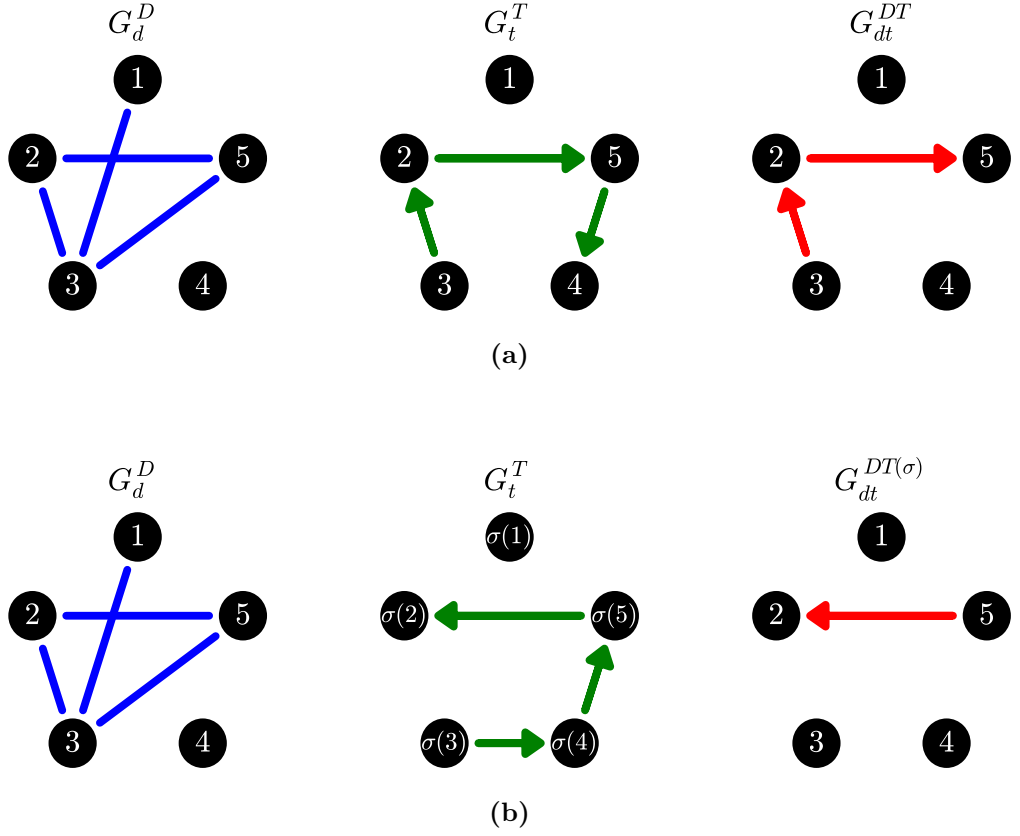


Figure 2.2: The relationship, for a simple set of event data, between the spatial network G_d^D , temporal network G_t^T , and network of space-time pairs (event network) G_{dt}^{DT} . The links of G_{dt}^{DT} are those where an link exists in that position in both G_d^D and G_t^T . The first panel a) shows the relationship for the original data, whereas b) shows how this is changed under a permutation σ (where $2 \leftrightarrow 4$) of the temporal data, as would be used in a Knox test using a Monte-Carlo approach.

2.3.4 Relationship with the Knox test

The derivation of an event network does not, in itself, afford any greater insight than was already possible via simple pair-wise comparison. Indeed, such a network is simply an alternative means of structuring the information concerning the closeness of events in both space and time, and the main statistics of interest can easily be expressed in terms of the network. The total number of close spatio-temporal pairs, for example, is equal to the total link-count of G_{dt}^{DT} , and other measurements can also be found by consulting the network.

A corollary of this is that existing pair-based tests of clustering, such as the Knox test, are also expressible in these terms. Although the translation is almost trivial, it is nevertheless instructive to briefly sketch it, since it provides a useful reference point for some of the arguments which follow. In addition to this, formulation of the Knox test in graphical terms is also illuminating as a means of understanding the test itself.

For a Knox test in which a close pair is defined by a spatial radius D and temporal radius T , the networks G_d^D , G_t^T and G_{dt}^{DT} can be built as described in the previous section. As noted previously, the first step in the analysis - to count the total number of close pairs - is achieved by simply counting the number of links in the network G_{dt}^{DT} , which itself is the intersection of G_d^D and G_t^T .

The substance of the test lies in the generation of a reference distribution via a permutation approach, in which the temporal data are repeatedly shuffled. Such a shuffling can be understood in network terms as a permutation of the vertices of G_t^T ; denoting the network of close space-time pairs under a permutation σ as $G_{dt}^{DT(\sigma)}$, the link-set of such a network is defined as:

$$E_{dt}^{DT(\sigma)} = \{(i, j) \mid (i, j) \in E_d^D \quad \textbf{and} \quad (\sigma(i), \sigma(j)) \in E_t^T\}. \quad (2.6)$$

The process can be understood visually by considering some physical layout of the networks. If each vertex is given some position (which is abstract, and does not correspond to the physical location of the event it represents), then it is immediately seen that the links of G_{dt}^{DT} are precisely those for which a link is present in the same position for both G_d^D and G_t^T . The result of a permutation σ can be understood by adapting this: if the positions of the vertices of G_t^T are shuffled (so that $\sigma(i)$ appears in the position of i), then $G_{dt}^{DT(\sigma)}$ can be constructed by identifying overlapping links. An example of this is shown in Figure 2.2b, demonstrating how the effect of the permutation is manifested in $G_{dt}^{DT(\sigma)}$.

In these terms, the Knox test is therefore carried out by finding the network $G_{dt}^{DT(\sigma)}$ for a number of choices of σ , and counting the links in each case. Significance and effect size are then computed exactly as previously. This is, of course, a trivial translation of the Knox test; however, the process can be extended by considering more complex features of the event networks.

2.3.5 Network motifs

Motif analysis is an approach to the study of networks which examines the occurrence of small subgraphs (containing, typically, only 3 or 4 vertices) within much larger networks. For a given network, the observed frequencies of various subgraphs are compared against those which would be expected under an appropriate null model, and this can be used to identify those subgraphs which occur disproportionately. These are referred to as *motifs*: small, atomic elements which characterise the network as a whole. Their over- or under-representation can be used to make inferences relating to, for example, the processes driving network formation.

To be clear, a motif of order n is a connected set of n vertices which has a particular configuration of links existing between its members. Since every possible configuration of links corresponds to a potential motif, numerous subgraphs are considered during analysis and the motifs identified are specific single examples. The analysis is easier to interpret, therefore, for small n , since the number of potential motifs is relatively small: there are, for example, only 13 connected directed networks of 3 vertices, up to isomorphism (see Milo *et al.*, 2002), reducing to 2 in the undirected case.

The number of possible motifs is reduced still further when considering event networks, due to the structural constraints imposed by the nature of the underlying data. The direction of network links, for example, is determined by the temporal precedence of events: since there is a temporal ordering in the data (which is transitive) any event network must therefore be cycle-free (trivially, for example, two

vertices can be connected in at most one direction). An event network is therefore an example of a directed acyclic graph (Bang-Jensen & Gutin, 2007). Restrictions such as these mean that only 24 motifs of order 4 (shown in Figure 2.3), and 4 motifs of order 3 (see Figure 2.4a), are possible for event networks.

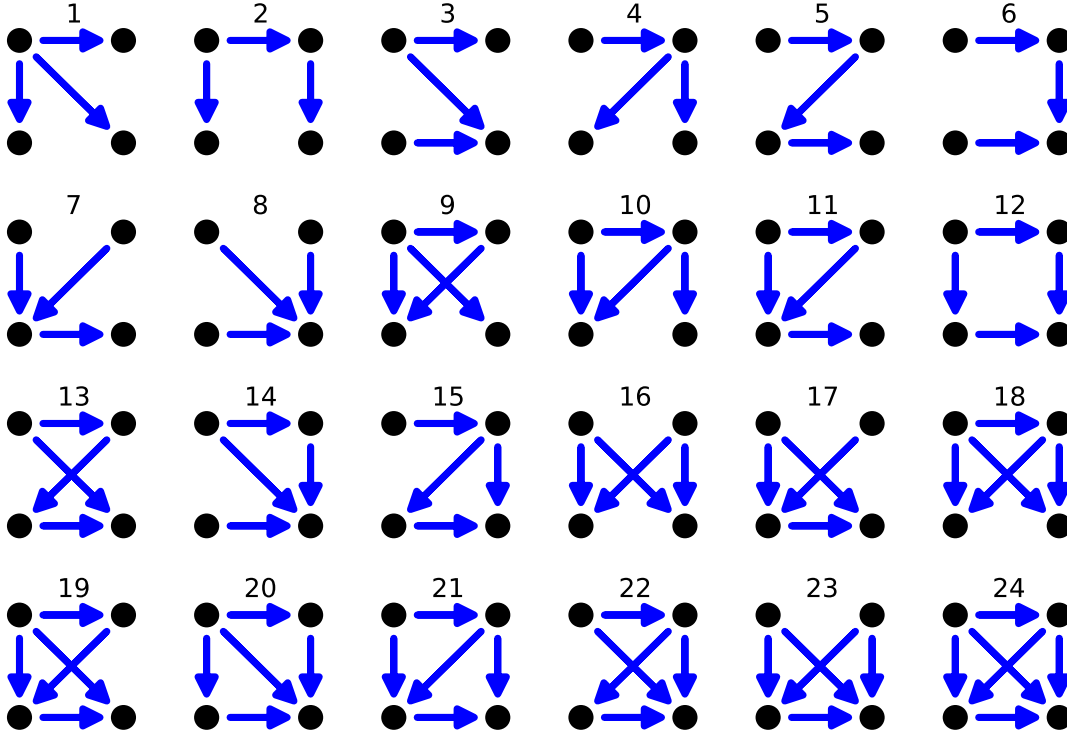


Figure 2.3: All possible 4-vertex motifs, up to isomorphism, which can occur in event networks. All subgraphs are cycle-free, as is necessary because of the temporal ordering of event data.

2.3.6 Motif interpretation

The value of motif analysis in the context of event data can best be understood by translating individual motifs back into the terms of the spatio-temporal patterns they represent. Each motif corresponds to a particular spatial configuration and temporal precedence among a small set of events, and it is the prevalence of these patterns which the analysis seeks ultimately to establish.

Although the meaning of each motif can be understood intuitively by simple visual inspection, the situations to which they correspond can also be expressed in concrete terms. As a demonstration of this, the generic patterns to which each of

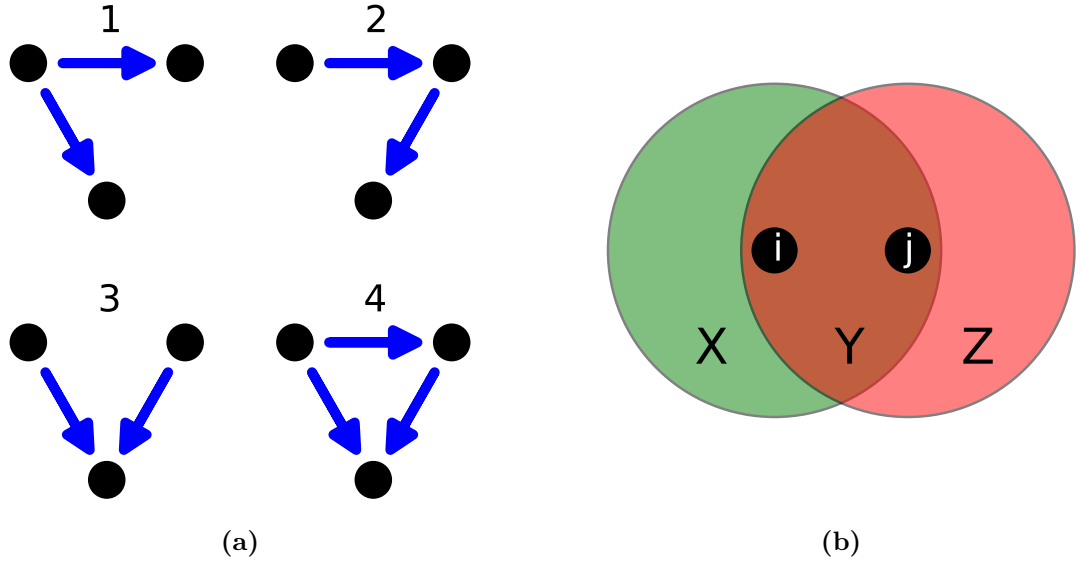


Figure 2.4: Motifs of 3 vertices and their real-world interpretation: a) all 3-vertex subgraphs which can arise in event networks, up to isomorphism, and b) two example events, i and j , with a circular region of radius D indicated for each.

the possible 3-vertex motifs shown in Figure 2.4a correspond are described in terms of hypothetical sets of events.

Each of the 3-vertex motifs contains at least one link, representing a close pair of events, and so the existence of two such events is taken as a common point of reference. For concreteness, Figure 2.4b shows two possible events, i and j , with regions of diameter D drawn around each, and the locations of events are described by reference to the various regions of this diagram. Without loss of generality, it is assumed that event i occurred prior to j .

Motif 1 The initial event i is followed by two other events within a temporal window T : event j , and another event occurring in the region **X**.

Motif 2 The initial event i is followed by another event, j , within time T . Event j is then followed by a third event, which **either**: a) also occurs within time T of i and lies in region **Z**, or b) occurs more than T after i and lies in either **Y** or **Z**.

Motif 3 Event i is followed by a close pair j , which was also preceded by a third

event occurring in the region \mathbf{Z} . Equivalently, two events occurred within time T but at a separation greater than D , but were followed by a third event occurring within D of both.

Motif 4 A close pair of events i and j are followed by a third event occurring in region \mathbf{Y} within time T of the first event i .

Similar reasoning can be used to characterise the patterns represented by the 4-motifs; however, full enumeration would be excessive and is not done here. As with the 3-vertex case, each motif can be understood intuitively as a distinctive pattern of events.

2.3.7 Event chains

The other network features to be analysed for event networks are chains, which are larger in terms of size, but conceptually simpler, than motifs. In this context, a *chain* is a sequence of events in which each can be linked to the previous one as a close pair. It is therefore a set of events for which a conceptual link runs throughout, via the antecedence relationship, even though only sequential pairs are connected.

In terms of the specific network features to which it refers, the term ‘chain’ can be defined in various ways, and so it is necessary to specify the sense in which it is used here before discussing the concept further. The definition is made with a number of theoretical concerns in mind, which will be discussed subsequently.

In an event network, a *chain* is a directed path of maximal length (*i.e.* one which is not a subset of a longer path).

Several aspects of this definition are worthy of comment. Firstly, the fact that chains are directed means that they respect the temporal precedence of events; in any chain, the events in question must have happened one after the other, in a definite ordering. It may be the case that larger groups could be found if directionality was disregarded (in network terms, these would be ‘weakly connected components’),

but these cannot be characterised easily as stylised series of events.

The ‘maximal’ aspect of the definition also has a number of implications. A path in an event network is only a chain if it cannot be extended further; in other words, the first vertex in the chain must have no in-links, and the final vertex must have no out-links. A chain of length 2, for example, does not comprise 2 chains of length 1, since this condition is violated in both cases. This also means that the number of chains of length 1 is not simply equal to the number of links in the network: only links which are isolated (*i.e.* not part of a longer chain) are counted as such.

This maximality has consequences in terms of what can be inferred from the identification of a chain. The existence of a chain of length l indicates not only the presence of a sequence of l events, but also the absence of any $(l + 1)$ th event with which the sequence could be extended. Chain length therefore corresponds to an upper limit on the linkage of events, in this sense, and suggests an alternative perspective: disproportionate incidence of chains of length l could be interpreted as evidence that $(l + 1)$ th events tend *not* to occur.

One final point concerns the uniqueness of chains, and the way in which they are counted during analysis. According to the definition, two chains are distinct if they differ in at least one link, and chain counts are calculated on the basis of unique chains. A given link may therefore feature in more than one unique chain, as long as they diverge at some stage. This is an unavoidable consequence when chains are defined in this way, and can be considered as both a positive and negative feature. In one sense, it leads to ‘double counting’ of certain paths; however, it could also be argued that paths which diverge at some point are genuinely distinct and should be treated as such.

Several of these issues are summarised by the hypothetical example shown in Figure 2.5. The diagram shows the 4 unique chains which are present in a hypothetical

6-vertex event network, together with their lengths. The effect of the directionality is manifested in the fact that the longest chains are of length 2, and the fact that those 2-chains are not broken into shorter fragments is due to the requirement of maximality.

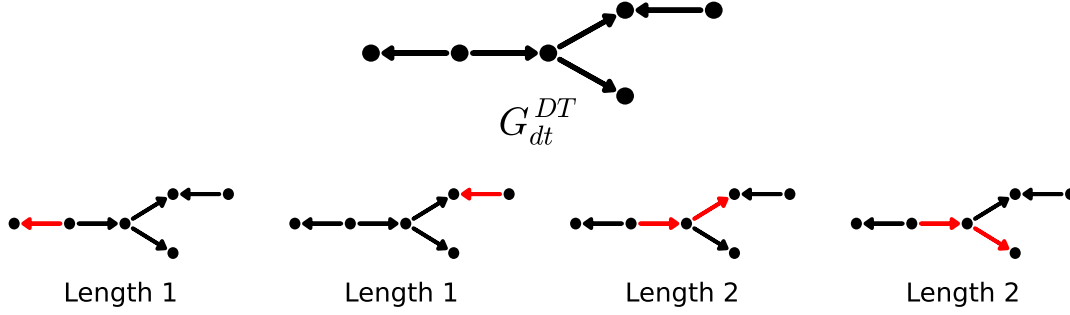


Figure 2.5: The identification of chains. The 6-vertex event network denoted by G_{dt}^{DT} contains four unique chains, each of which is shown below. In each case, the chain in question is indicated by red links and its length is stated.

The interpretation of chains, in terms of the real-world patterns which they represent, is fairly straightforward. Assuming that the close pair relationship can be taken as evidence of conceptual linkage, analysis of chain length can be interpreted as an indication of the typical size of ‘bursts’ of activity. It represents a measure of the extent to which events tend to occur in ‘spates’ of sustained activity, and also provides a means of examining the extent to which the clustering of crime is fragmentary.

Extending this further, such results can also be used to make inferences about the behaviour of offenders. Sequences of close events are often hypothesised to be the work of the same offender, with arguments related to the concept of foraging frequently invoked (*e.g.* Johnson *et al.*, 2009b). In such a context, the length of chains acts as a measure of the extent to which offenders engage in such bursts of behaviour before moving on. The decision to end such a sequence, and the stage at which it is taken, is a theoretically-important one, and relates to notions of ‘carrying capacity’ when crime is modelled by analogy with ecology.

Chains are not a feature for which spatial patterns can be characterised with any degree of certainty. The existence of a chain is consistent with ‘drifting’ behaviour (where locations move in some definite direction as the chain progresses) but also with repeated occurrences at a fixed location. Little can therefore be inferred about spatial trends, and the primary value is in the inferences which can be made about the ‘burstiness’ of offending patterns.

2.3.8 Statistical analysis

Having introduced the concept of network motifs and chains, and the interpretation of such features in a spatio-temporal context, it remains to discuss the methods by which the prevalence of such subgraphs can be evaluated. To emphasise again, the motivation for this analysis is to gain additional insight into sets of events which are known to be clustered (this having been demonstrated via an established statistical test, such as the Knox test) but where the exact nature of the clustering is unknown. Several factors specific to the case of event networks mean that established techniques cannot be applied without modification.

The process of counting subgraphs of small order in a given network, on which this analysis rests, is well-documented. Brute-force enumeration can be used for sufficiently small networks (and is used in the work in this chapter) but, where this is computationally prohibitive, efficient methods based on sampling are also available (Kashtan *et al.*, 2004; Wernicke & Rasche, 2006). The primary technical question shifts, therefore, to that of which ‘random’ networks the observed data should be compared against. Crucially, these networks should correspond to situations for which the level of clustering is the same, in a purely pair-wise sense, as for the original data, in order to gain insight beyond that which can be found by existing means. Since the method is intended to be applied to data for which clustering is already known to be present, this is a crucial point.

For the generation of networks for comparison, one possibility would be to em-

ulate the approach of the Knox test, described in Section 2.3.2, by employing a permutation approach. Such an approach, however, is immediately seen to be problematic upon consideration of the relationship between link count and subgraph frequency. It is a simple observation that, for networks with a constant number of vertices, there will exist an association between the number of links and the frequency of subgraphs; to choose two pathological cases, an empty network contains no connected subgraphs, whereas every subgraph of a complete network is itself complete. Away from these extremes, an increasing density of links will generally imply a higher number of connected subgraphs overall (ultimately favouring more dense subgraphs and longer chains).

The issue is, therefore, that the event network of any set of clustered events (if the definition of ‘clustered’ is taken to be ‘that which gives a significant Knox test’) will necessarily have a lower link count when the temporal data are permuted, since that is precisely what is indicated by the Knox test. Any difference between observed subgraph counts and those under permutation might, therefore, simply be an artefact of the change in density, rather than any effect over and above the known clustering. Since observed subgraph counts are *expected* to be anomalous under such a comparison, they are therefore of little inferential value.

The implication of this argument, from an analytical perspective, is that it is necessary to maintain a constant link-count when generating random networks against which observed subgraph frequencies can be compared. This is certainly not a novel observation for motif analysis, and methods by which the issue can be addressed are outlined in previous work on the topic (*e.g.* Milo *et al.*, 2002). The typical approach taken is to generate the required randomised networks by simply re-wiring the links of the original observed network; that is, by randomly re-assigning one or both of the end-points of individual links. In this way, no links are created or destroyed, and other structural features of the network (*e.g.* vertex degrees) can also be preserved, as in the configuration model (Newman *et al.*, 2001). It is assumed that networks

derived by such a method are a representative random sample of all networks which possess the prescribed properties.

This approach, however, is not applicable in the context of event networks. As noted previously, the fact that event networks are derived from spatial and temporal data constrains the space of possible configurations: they must, for example, be cycle-free (since otherwise this would imply that one event occurred simultaneously before and after another).

In addition, since links represent spatial proximity, their existence is not mutually independent, and there exist certain combinations which cannot arise from any possible set of events. To explain this, it is useful to return to the concept of the unit disk graph (UDG), of which the spatial proximity network G_d^D is an example. A *realisation* of a given UDG, G' , is a set of points in space for which G' is the induced UDG; that is, a set of points whose proximity is as indicated by G' . Any UDG must have at least one realisation, by definition. A key problem in the field therefore concerns the question of whether a given arbitrary graph has a realisation, *i.e.* whether it is a UDG.

Not all graphs do have a realisation: as shown by Marathe *et al.* (1995), for example, the 7-vertex ‘star’ graph shown in Figure 2.6 does not. This is because, in 2-dimensional space, it is impossible for six circles of equal radius to intersect with another similar circle (the red circle in Figure 2.6b) without at least two of the six intersecting with each other. Aside from stylised examples such as this, however, establishing the (non-)existence of a realisation for arbitrary graphs is non-trivial; indeed, it has been shown by Breu & Kirkpatrick (1998) to be NP-hard.

The implication of these results for the present study is that some networks do not arise as the event network of any possible set of events. Motivated by this, it is therefore helpful to define a *valid* network as one which is the event network for

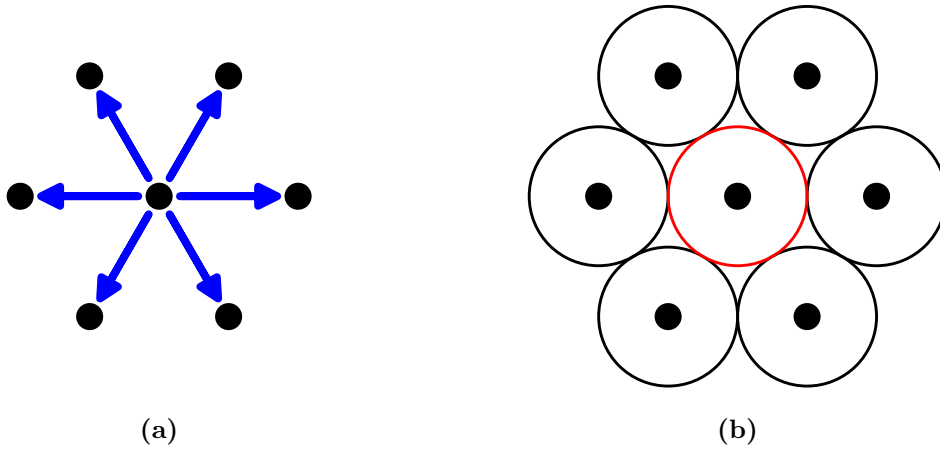


Figure 2.6: A ‘star’ graph of 7 vertices, which cannot arise as an event network due to geometric constraints. Were the graph shown in a) to be an event network, it would imply that the events represented by the outer vertices all occurred within distance D of that represented by the central vertex. However, by plotting circles of diameter D around each location in b), it is clear to see that, if the outer circles intersect with the red central circle as required, at least two of the outer circles must also intersect. The corresponding locations must therefore lie within D of each other, meaning that an additional link must exist.

at least one set of spatio-temporal event data. Only networks which are realisable in this sense must be considered in analysis, but it is clear that: a) not all networks are valid, and b) determination of the validity of a given network is, in general, computationally-prohibitive. It is for this reason that a simple re-wiring approach is inappropriate: it is liable to generate invalid networks, and there is no natural way to adapt the process to ensure validity.

An even more significant, though more subtle, problem concerns the fact that, even amongst valid networks, not all are equally likely to arise from random event data. The mapping from event datasets to their event networks is many-to-one, and the number of datasets which give rise to each event network is not equal; some arrangements will naturally arise more readily than others. Given that the ultimate interest of the analysis is to establish the randomness (or departure from it) of the event data, it is evident that, even if the set of all event networks could be characterised succinctly, randomly sampling from it would still be inadequate.

2.3.9 Random network generation

In light of these observations, it is clear that the generation of appropriate randomised networks cannot be done while remaining agnostic to the nature of the underlying data, and that a novel method is therefore required. Such a method should produce networks which are valid event networks (*i.e.* realisable from some set of spatio-temporal events) and which have the same link-count as the event network induced from the observed data.

It is also necessary to impose one further requirement. Events can be clustered in three senses: in space, in time, and simultaneously in both. While the ultimate objective here is to characterise the latter of these, it is influenced by the former two: all else being equal, a more spatially-clustered dataset will tend to give rise to more close pairs in space and time, for example. To control for this, generated networks should also correspond to sets of events which are as clustered as the observed data in both space and time independently. That is, their spatial clustering should be equivalent to that of the observed data when time is ignored, and *vice versa*.

The solution proposed here is to explicitly simulate synthetic sets of events in time and space, constructed so as to have the required level of clustering in each dimension. Because events are simulated directly, validity is immediate, so that the substance of the task is concentrated on achieving the correct level of clustering.

A trivial way to ensure the required levels of spatial and temporal clustering would be to simply preserve the spatial and temporal distributions exactly, so that the locations and times of synthetic events are exactly those of the observed data, but shuffled. This is the approach taken by the permutation version of the Knox test, as described in Section 2.3.2. In the context of random network generation, the task would be to do this in such a way that the number of space-time pairs was as required, *i.e.* equal to that of the observed data. Doing this while maintaining the exact spatial and temporal distributions is, however, a very strict limitation, and

such an approach is infeasible: generating a synthetic network in this way would be equivalent to seeking a permutation of the temporal data which gives exactly the same number of close space-time pairs as for the un-permuted data. This is not guaranteed to exist.

In order to make progress, therefore, the sense in which the spatial and temporal clustering should be matched is relaxed somewhat. Rather than insisting that the distributions are exactly equal, it is required only that the set of events produced should have the same number of close spatial pairs and same number of close temporal pairs as the observed data (as well as, of course, the same number of space-time pairs). Though the exact distributions may be different, this does mean that the events will be ‘as clustered’ in each dimension, in at least a basic sense, as the observed data. Defining the spatial, temporal and spatio-temporal networks derived from the synthetic data as \tilde{G}_d^D , \tilde{G}_t^T and \tilde{G}_{dt}^{DT} respectively, the requirement is that each of these should have the same link-count as its counterpart for the observed data.

To recapitulate the analytical situation, the objective of the process is to produce, given a set of N events occurring within some spatially- and temporally-limited region, another set of synthetic events with the same number of spatial, temporal and spatio-temporal close pairs. These values, which are equal to the link-counts of the corresponding networks, will be denoted by M_d , M_t and M_{dt} for the observed events (dropping superscripts without loss of generality).

The production of the synthetic events is achieved by taking a set of fully random events (in space and time) as a start point, and making iterative adjustments to it until the required conditions are satisfied. The process is initialised by generating N events within the region in question, uniformly at random in both space and time. The number of close pairs in space, time, and both space and time are calculated for these events, and denoted \tilde{M}_d , \tilde{M}_t and \tilde{M}_{dt} respectively. An ‘energy’,

E , is used in order to quantify the extent to which these measurements differ from those required:

$$E = \frac{|\widetilde{M}_d - M_d|}{\widetilde{M}_d + M_d} + \frac{|\widetilde{M}_t - M_t|}{\widetilde{M}_t + M_t} + \frac{|\widetilde{M}_{dt} - M_{dt}|}{\widetilde{M}_{dt} + M_{dt}} \quad (2.7)$$

This energy is zero only when the required counts are equal for the observed and synthetic data, and increases as the deviation increases. The process which follows therefore makes repeated changes to the event set with the aim of reducing E until it is zero.

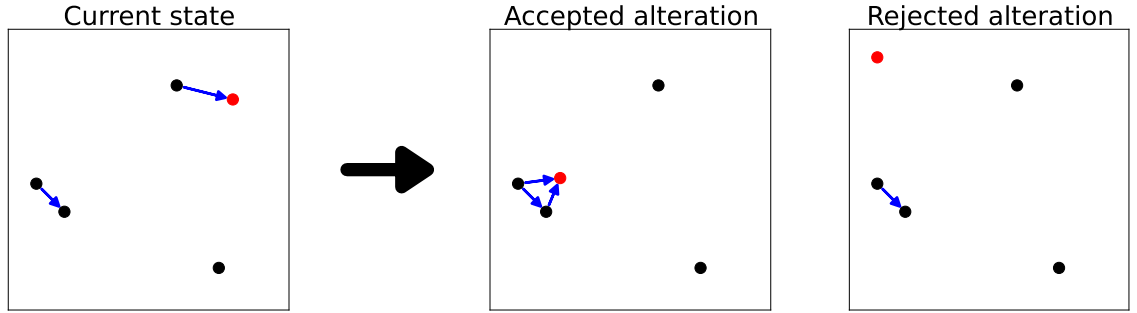


Figure 2.7: An illustrative example of the process of moving points so as to find a configuration with the correct level of spatio-temporal clustering. The objective, in this hypothetical case, is to find a configuration with 4 spatio-temporal close pairs and, for simplicity, it is assumed that all events shown are close temporally. The left panel shows the configuration at the start of a single iteration, in which 2 close pairs are present. The point highlighted in red is selected for movement: one possible change leads to a distribution with 3 close pairs and would be retained; the other reduces the number of pairs to 1 and would be discarded.

Each iteration involves first establishing the current energy, E^{cur} , and then selecting one of the events, uniformly at random, as a candidate for adjustment. A prospective new position and time are then generated for the selected event, chosen uniformly at random in the known spatio-temporal region. The value of the energy under this hypothetical change, E^{new} , is calculated and compared with E^{cur} . If $E^{new} < E^{cur}$, the prospective change is made definite; otherwise, it is discarded and the original positions are retained (see Figure 2.7). In this way, E decreases monotonically as the algorithm iterates, and the process eventually ends when it reaches zero. A small random error can be introduced whereby energy-increasing changes are accepted (to avoid local minima, in the spirit of simulated annealing - see Newman & Barkema,

1999) but this is not found to be necessary in the practical examples considered in Section 2.4.

Expressed informally, the process involves the repeated random movement of individual points (in space and time), with the objective at each stage of moving closer to the desired configuration. Movements which do this are retained, while others are not; in this way, a suitable configuration ultimately is found.

Using this process, it is therefore possible - given a set of events, a spatial threshold D and a temporal threshold T - to produce synthetic sets of events with the same level of clustering in either dimension, and the same number of close pairs in both space and time. As such, they are matched on the type of pair-wise clustering measured by techniques like the Knox test, whilst being random in all other respects. Because of this, the occurrence of other features, such as motifs and chains, can be meaningfully compared between observed and synthetic data, in a way which controls for basic pair-wise clustering.

2.4 Results

The purpose of the techniques described in the previous section is to provide a means by which patterns of space-time clustering can be characterised and differentiated. The central idea is that, by expressing inter-event relationships as a network, it is possible to identify more complex patterns, and that this can be done in a statistically-rigorous way. This section demonstrates the application of these techniques to real-world datasets relating to criminal events; the study of crime is a field in which such clustering is of theoretical importance, and was the motivating case for this work. The data considered cover two crime types, residential burglary and maritime piracy, both of which have been shown to cluster in space and time (Johnson *et al.*, 2007; Marchione & Johnson, 2013), but which are of markedly different character in many other respects.

2.4.1 Data

The burglary events considered consist of 5,690 events, representing all residential burglaries which took place in the city of Birmingham, UK, between March 2012 and February 2013, inclusive. For each event, the location of the victimised property is recorded in terms of a British National Grid co-ordinate reference given by police, to an accuracy of 1 metre. The temporal information is given in the form of a window, representing the earliest and latest possible times at which the event could have taken place. In order to establish a point estimate of an event’s time, the midpoint of these times is taken in all cases (other approaches can be used when timing is uncertain - see Section 4.3.1 - but that is not the focus at this stage).

The piracy events used in the analysis are the 545 pirate attacks recorded by International Maritime Organisation as occurring during the year 2010. For each of these events, the best known location is provided in terms of a latitude/longitude grid reference, and the day on which the attack occurred is also given. These are sufficient to allow the required pair-wise separations between events to be calculated in each case.

2.4.2 Comparing datasets

One issue to be considered when comparing the analysis for these two datasets in tandem is the disparity in spatial scales between the two cases. The burglaries, for example, are contained entirely within an urban area, and analysis of this crime typically defines close pairs as those which have occurred within a few hundred metres of each other (Johnson *et al.*, 2007). Pirate attacks, on the other hand, are distributed at a global scale, and a distance of less than 100 kilometres might be regarded as ‘close’ (Marchione & Johnson, 2013).

The distances involved have different implications in either case: variation in terrain and mode of transport mean that the values are not directly comparable. Indeed, the practical implications (which this research seeks ultimately to inform) are also

rather different: seas are sparse environments in which policing typically involves the monitoring of large areas, whereas interventions against burglary would usually be implemented at a much more localised level.

This is an important consideration, since the results of any clustering analysis of this form depend crucially on the choice of spatial and temporal thresholds which define a ‘close pair’. These determine the amount of information encoded in an event network: if the thresholds are too small, the sparsity of links means that few details can be extracted; if they are too high, however, too many links exist, the concept of closeness is diluted, and meaningful patterns might be concealed. Indeed, if analysis is carried out for several scales, the range over which significant results are found (if any) can be interpreted as defining a typical scale for clustering.

In order to compare results between the datasets, therefore, it is necessary to establish some equivalence between their scales from the perspective of clustering. This can be done by considering the number of close space-time pairs in the dataset, relative to the total number of events (in terms of the network G_{dt}^{DT} , this is equal to the number of links divided by the number of vertices, which is half of the mean degree). This gives an approximate measure of the concentration of links present in the network (the usual network measure of *density* is not used, since its denominator - the maximum number of possible links - is not well-defined for event networks). The value of this ratio is given for a variety of spatial and temporal thresholds, for both event types, in Table 2.1.

While there is substantial variation across the thresholds considered, Tables 2.1a and 2.1b contain values of broadly similar order. This suggests that the threshold values shown give rise to networks of approximately equivalent density in each case, and that they can therefore be compared meaningfully. Moreover, the presence of near-equal values in both tables - for example, the values 1.29 for (21 days, 400m) in the burglary case and 1.25 for (21 days, 200km) in piracy, which are indicated

	100m	200	300	400	500		50km	100	150	200	250
7 days	0.10	0.21	0.36	0.54	0.74	7 days	0.29	0.39	0.52	0.61	0.72
14	0.15	0.33	0.60	0.92	1.29	14	0.43	0.57	0.78	0.90	1.08
21	0.18	0.44	0.82	1.29	1.83	21	0.57	0.77	1.06	1.25	1.50
28	0.22	0.56	1.05	1.66	2.36	28	0.66	0.90	1.25	1.46	1.79
35	0.26	0.68	1.27	2.02	2.87	35	0.70	0.99	1.41	1.66	2.07

(a) Residential burglary

(b) Maritime piracy

Table 2.1: The ratio of the number of close space-time pairs to the total number of events, for various spatial and temporal thresholds. The dashed lines indicate approximately equal values; the corresponding threshold combinations are used to compare results between crimes.

by dashed lines - invites direct comparison between the corresponding cases. These will be considered alongside the general patterns across all possible combinations.

2.4.3 Traditional clustering analysis

As stated previously, the primary motivation for the approach described in this chapter is to offer additional insight into the structure of spatio-temporal datasets which are already known to be clustered in a pair-wise sense. In order to formally establish this base, therefore, a brief examination of the pair-wise clustering present in the two datasets was performed.

The clustering was measured by means of a traditional Knox-type permutation test, as described in Section 2.3.2. In this case, the Monte Carlo simulation consisted of $n_K = 99$ iterations, each of which involved random permutation of the temporal data for the events. Pseudo-significance was calculated by applying equation (2.2), and z -scores used to quantify the magnitude of deviation from the null distribution. These values are shown, again for several possible threshold values, in Table 2.2; the clustering is significant at the 0.01 confidence level in each case.

The results confirm the expected result, that both sets of events exhibit highly significant clustering in space and time, consistently across several definitions of the

	100m	200	300	400	500		50km	100	150	200	250
7 days	16.00	12.92	12.60	12.61	14.25	7 days	13.51	12.88	13.82	13.65	13.84
14	13.03	11.76	9.24	9.99	9.33	14	12.45	10.71	10.68	11.89	11.99
21	10.61	9.90	7.91	9.25	9.01	21	11.97	10.46	11.10	12.53	10.92
28	8.64	9.17	9.65	8.00	9.91	28	11.23	9.88	9.93	9.86	9.33
35	9.98	10.32	8.70	9.56	9.25	35	9.42	6.47	9.55	7.93	7.94

(a) Residential burglary

(b) Maritime piracy

Table 2.2: Knox analysis: values shown are z -scores for the observed number of close pairs, relative to those found under the randomised distribution. In all cases the observed value was higher than that for all randomised datasets, corresponding to a pseudo-significance of 0.01.

space-time window. As expected, the magnitude of the effect decreases as granularity is lost in the data, suggesting that the effects are distinctly local in both space and time. The effect appears to be slightly larger in the case of piracy, though the disparity between crimes is not particularly large; certainly the values are consistent with those published elsewhere (Johnson *et al.*, 2007; Marchione & Johnson, 2013).

2.4.4 Identification of motifs

The fact that the two crimes considered exhibit space-time clustering means that both datasets are of the type to which the techniques proposed in this chapter are intended to apply. This section therefore describes that application, considering first the occurrence of 3- and 4-vertex motifs in the induced event networks. In order to do this while controlling for the pair-wise clustering which was established in the previous section, observed counts were compared against those measured for synthetic datasets, generated as described in Section 2.3.9. Since these were constructed so as to have the same level of pair-wise clustering as the observed events, any significant patterns observed must therefore be due to variation in the *type* of clustering present, rather than just its presence.

The motifs examined in the analysis comprised all possible 3- and 4-vertex sub-graphs which can occur in event networks, *i.e.* all those which were summarised in

Figures 2.3 and 2.4a. Particular motifs are referred to throughout by their indices in those figures, though small diagrams are also incorporated into the majority of figures.

As will be shown, the analysis was carried out for a range of definitions of the ‘close pair’ relationship. For any given spatial threshold D and temporal threshold T , the various networks described in Section 2.3.3 were first constructed: the spatial and temporal proximity networks, G_d^D and G_t^T , and the event network, G_{dt}^{DT} , which represents closeness in space and time. The occurrences in G_{dt}^{DT} of each possible motif were counted and recorded, as were the numbers of links in all three networks.

The next step was to produce 99 synthetic sets of events, using the process described in Section 2.3.9, for comparison. Each one of these was constructed so that each of the networks induced from it, \tilde{G}_d^D , \tilde{G}_t^T and \tilde{G}_{dt}^{DT} , had the same number of links as its observed counterpart (*i.e.* the events included the same number of close pairs in each dimension). For each of the 99 networks \tilde{G}_{dt}^{DT} , the number of occurrences of each motif was counted and recorded.

For each possible motif, therefore, 99 ‘null’ counts were produced, against which the observed count from G_{dt}^{DT} could be compared. This was done in the same manner as for the Knox test, by examining the extremity of the observed value relative to the null counts. For the pseudo-significance, the null counts were placed in an ordered list and the prospective rank of the observed value in this list, r_m , found; the p -value was then calculated as $\frac{r_m}{100}$. A p -value of 0.01 therefore indicates that the observed value was more extreme than any of the 99 reference values. In addition, and as before, the magnitude of the effect was measured by computing a z -score: the number of standard deviations by which the observed value differed from the mean of the null distribution.

Considering the 3-vertex case first, Figure 2.8 shows the results of this analysis,

for particular spatio-temporal thresholds, for each of the datasets. The thresholds chosen - (21 days, 400m) for burglary, and (21 days, 200km) for piracy - correspond to an approximately equal concentration of links within the networks (see Table 2.1), so that the comparison is an unbiased one, in this sense. The figures show, for each of the 4 possible 3-vertex motifs, box-plots of the distribution of counts within the synthetic networks, along with a marker for the observed frequency. For any particular motif, therefore, over- or under-representation in the data corresponds to the degree of deviation of the observed value from the box-plot area.

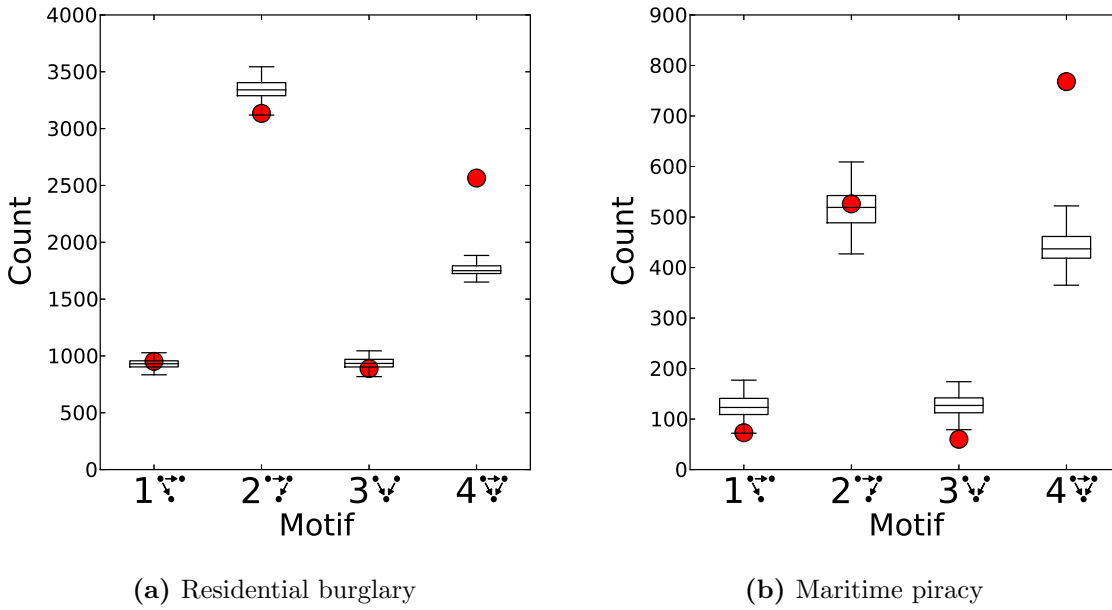


Figure 2.8: Comparison of observed and randomised counts of 3-vertex motifs for a) residential burglary, with a close pair threshold of 21 days and 400 metres, and b) maritime piracy, with a threshold of 21 days and 200 kilometres. Observed counts are indicated by circular markers, and the box-plots give the distribution of counts across the 99 simulated datasets in each case.

The most striking result in both cases is the highly significant over-representation of motif 4, which appears with a frequency far above what would be expected on the basis of chance. That this motif - a fully connected triple - should be so prevalent suggests that the occurrence of links in these networks is not independent, and that they have a significant tendency to co-occur in this way. In terms of the underlying crime events, the result suggests that events tend to occur not just in pairs, but in fact in larger groups within a restricted spatio-temporal domain.

The only other statistically significant result for these thresholds is the under-representation of motif 3 in the piracy data (though the magnitude of the effect is relatively small). This motif represents a situation in which an event occurs in the common neighbourhood of two previous events which were not themselves a close pair. Though several explanations are possible, the pattern could be interpreted as corresponding to a situation in which the two initial events are committed by different offenders; if the same offender was responsible, such a pattern would imply a reversal of direction. The relative absence of the pattern would therefore be consistent with the hypothesis that the presence of multiple active offenders in a given region is unlikely.

Aside from the identification of individual motifs, the results can also be interpreted as a whole. From this perspective, the near-total lack of significance for patterns other than motif 4 suggests that the majority of the fine-scale clustering present in the data can be attributed to configurations of that type.

The results shown in Figure 2.8 are for specific spatial and temporal thresholds, but the analysis can, of course, be carried out for arbitrary values. Figure 2.9 summarises the results for each of the spatio-temporal thresholds for which Knox values were given in Figure 2.2. For each possible motif, the effect sizes calculated for each combination of threshold values are represented by an array of colours.

Although it would certainly be excessive to consider each case in detail, several general trends are apparent. The significance of motif 4, for example, is ubiquitous across all thresholds, suggesting again that the dense clustering to which it corresponds is a fundamental signature of both processes. In the case of piracy, its effect size is noticeably larger at smaller spatial scales but larger temporal scales, which can be interpreted as implying a characteristic scale, in some sense, for behaviour of this type. In the case of burglary, the largest effects are seen at the shortest spatial

scales, which again may be instructive as to the time course of such victimisation.

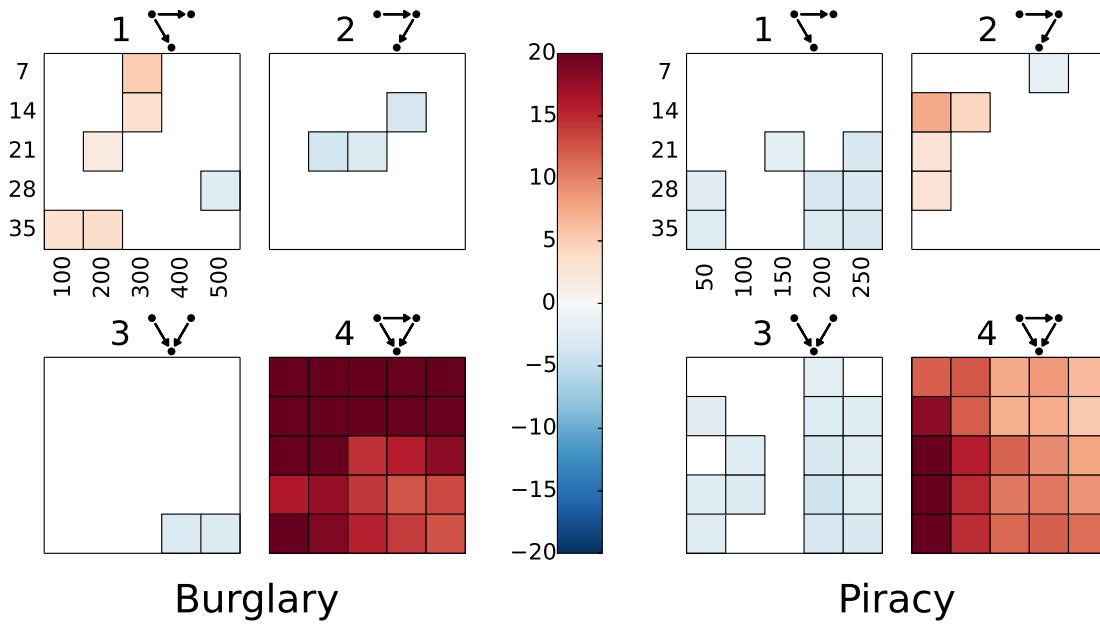
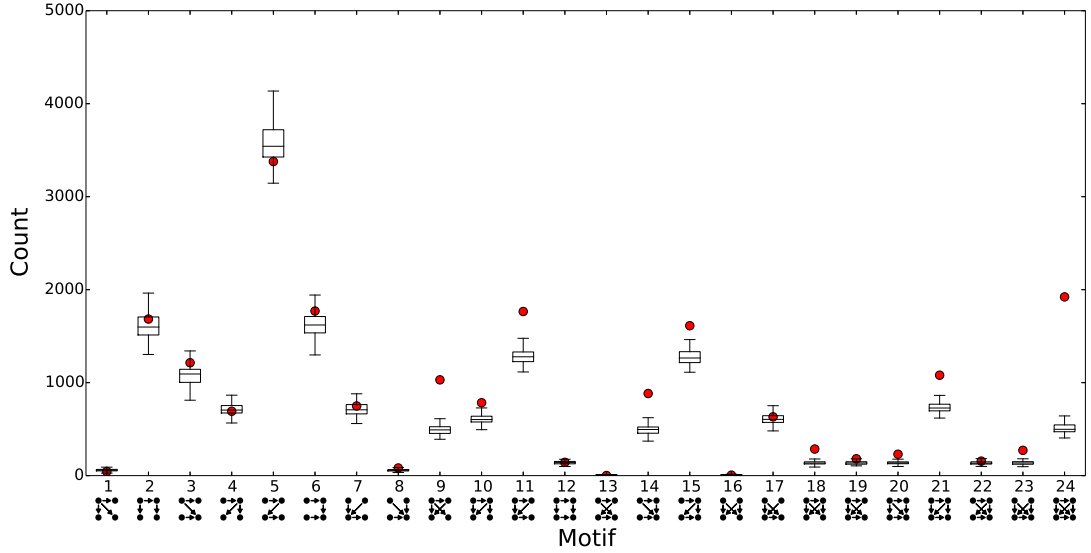


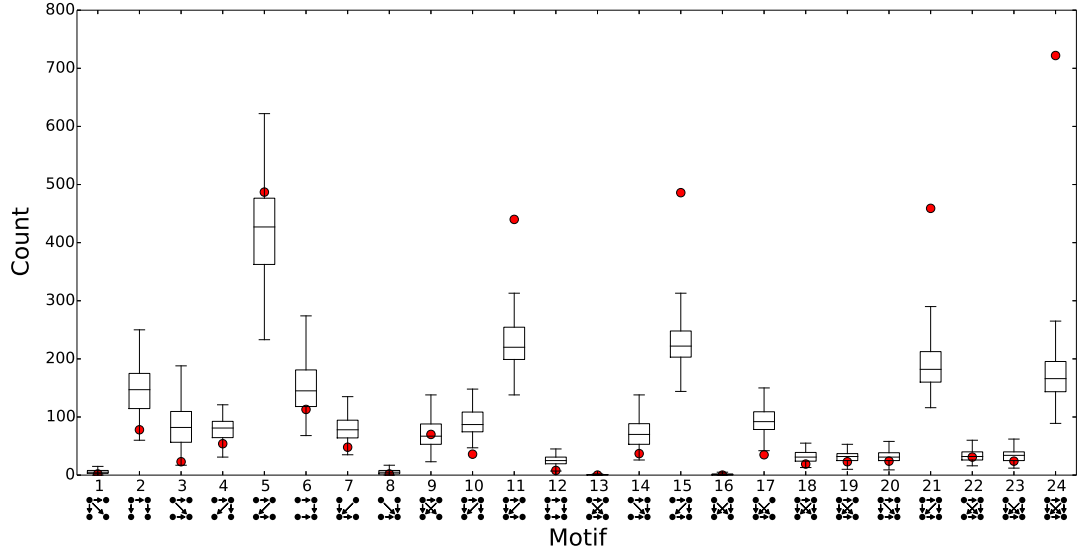
Figure 2.9: Statistical analysis of 3-vertex motifs for a variety of spatio-temporal thresholds. For any given cell, a black outline indicates that the frequency of the given motif is significant at 0.01 level; in such cases, the colour of the cell corresponds to the z -score for the effect, with blue indicating that the motif is under-represented and red that it is over-represented.

In addition, referring to Figure 2.9, piracy shows significant under-representation for motifs 1 and 3 in several cases. A hypothetical explanation can again be offered: such motifs correspond, in broad terms, to more spatially-diverse offending, and their relative absence might be suggestive of localised territorial behaviour by a relatively small number of offenders.

Motif 1 is also seen to be significantly over-represented for burglary in certain threshold combinations (though there is no clear trend as to which). Although it is impossible to say with certainty why this is the case, a theoretically-grounded option might invoke the concept of spatial ‘anchoring’ on the part of offenders: an offender who persistently victimises properties within some radius of a central anchor point - a ‘central place forager’, for example (see Johnson, 2014) - might be expected to generate such a pattern.



(a) Residential burglary



(b) Maritime piracy

Figure 2.10: Comparison of observed and simulated counts of 4-vertex motifs for a) residential burglary, with a close pair threshold of 21 days and 400 metres, and b) maritime piracy, with a threshold of 21 days and 200 kilometres. Observed counts are indicated by circular markers, and the box-plots give the distribution of counts across the 99 simulated datasets in each case.

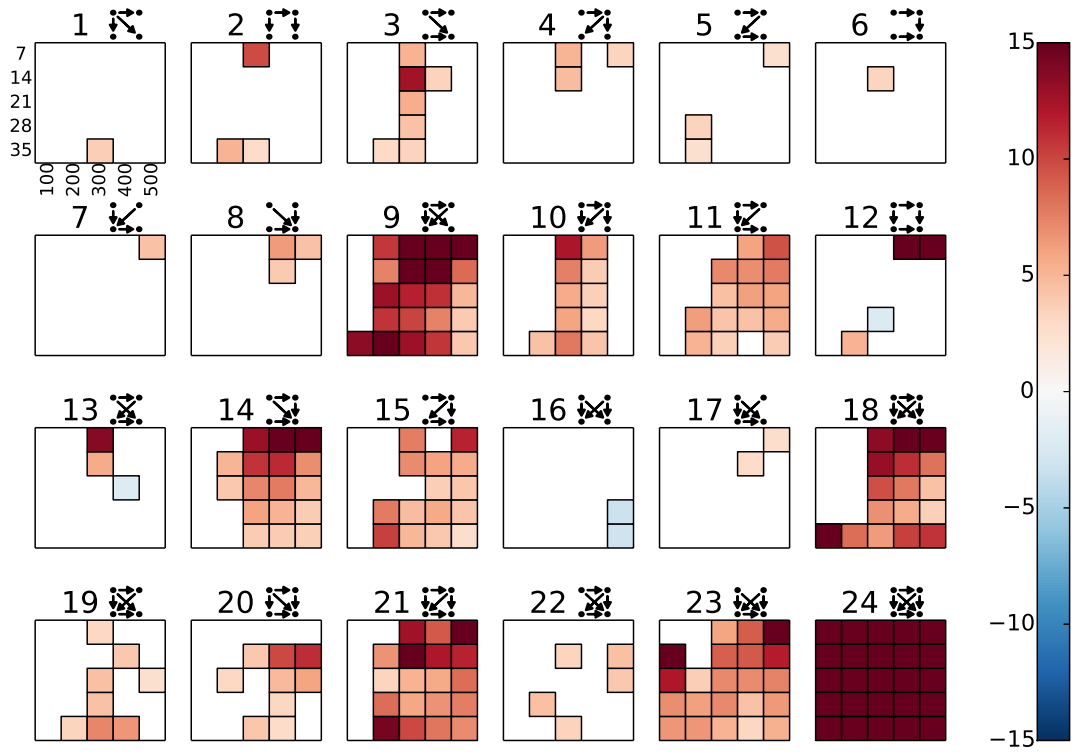
Changing focus to consider 4-vertex motifs, significant results can again be seen for both crime types. Figure 2.10 shows the comparison between observed and randomised counts for all such features, using the same spatial and temporal thresholds as for Figure 2.8. Again, several subgraphs are seen to occur disproportionately in each case.

In particular, motifs 11, 15, 21 and 24 show highly significant over-representation in both cases. Motif 24, the fully-connected pattern, is analogous to the fully-connected motif in the 3-vertex case, and its significance here constitutes further evidence of particularly dense clustering in both datasets. The reason for the appearance of each of the other three are not intuitively obvious, but all contain the fully-connected 3-vertex motif as a subgraph, augmented in various ways by an additional event.

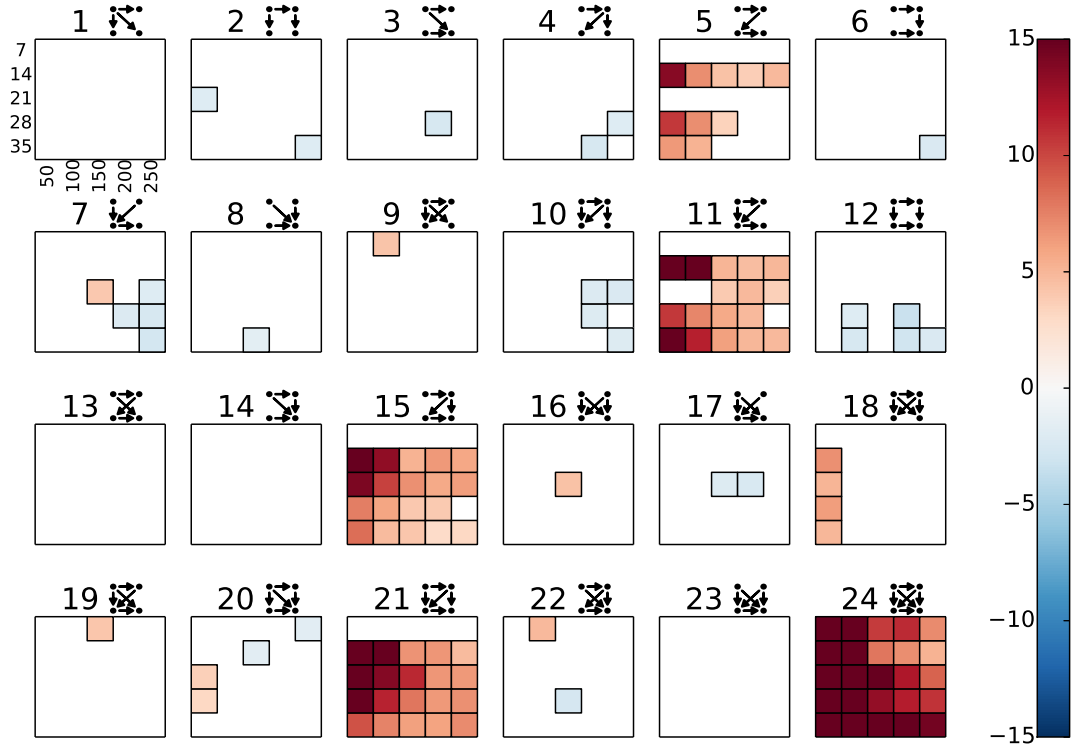
A summary of the results for various spatial and temporal thresholds is shown in Figure 2.11. The number of subgraphs under consideration (and their close similarity) means that there is little value in providing here an exhaustive exposition of each, though clearly there is scope for doing so. Rather than the presence of specific motifs, though, perhaps the most significant observation in this case is the difference between event types in terms of the patterns identified. Although the four motifs listed above are, for the majority of thresholds, significant in both cases, several motifs display contrasting behaviour.

Particularly clear examples, from visual inspection, are motifs 9, 10, 14, 18 and 23, which are highly significant in the case of burglary but barely so at all in the case of piracy. Indeed, several motifs are seen to be significantly under-represented in the case of piracy, such as 7, 10 and 12. Though the precise patterns to which they correspond may be intricate, those configurations do, in general, appear to represent spatially-diverse patterns of events, the unlikeliness of which can be reconciled with the maritime setting.

The higher overall prevalence of motifs in the case of burglary case also indicates a more dense, though still stylised, pattern of offending. Again, however, the most significant observation is simply the difference in patterns between the two event types, which suggests that the techniques employed here, and the study of motifs in particular, are capable of revealing differences in the patterning of offences at a deeper level than previously possible.



(a) Residential burglary



(b) Maritime piracy

Figure 2.11: Statistical analysis of 4-vertex motifs for a variety of spatio-temporal thresholds. For any given cell, a black outline indicates that the frequency of the given motif is significant at 0.01 level; in such cases, the colour of the cell corresponds to the z-score for the effect, with blue indicating that the motif is under-represented and red that it is over-represented.

2.4.5 Analysis of chain length

In addition to the analysis of motifs, the event networks were also examined for the occurrence of chains. This was measured statistically using exactly the same method as described in the previous section; indeed, the randomised networks were the same in both cases, with the only difference being the features measured.

Again, the results for particular combinations of spatial and temporal threshold can be presented as a boxplot, and this is done for both crimes in Figure 2.12. The threshold combinations used in these cases are (14 days, 500m) for burglary and (28 days, 150m) for piracy; for reasons of clarity, these differ from those used in previous similar figures, but are still matched in terms of the density of the networks (see Table 2.1).

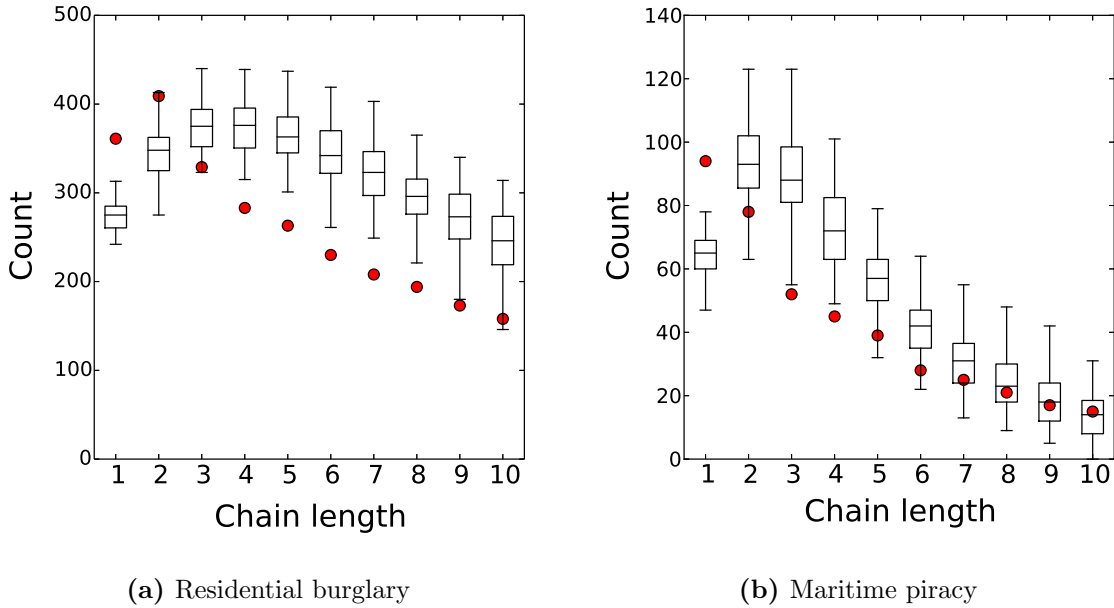


Figure 2.12: Comparison of observed and simulated counts of chains for a) residential burglary, with a close pair threshold of 14 days and 500 metres, and b) maritime piracy, with a threshold of 28 days and 150 kilometres. Observed counts are indicated by circular markers, and the box-plots give the distribution of counts across the 99 simulated datasets in each case.

Anomalous behaviour can be observed for both crimes at certain chain lengths. In particular, both crimes show an over-abundance of chains of length 1; that is, iso-

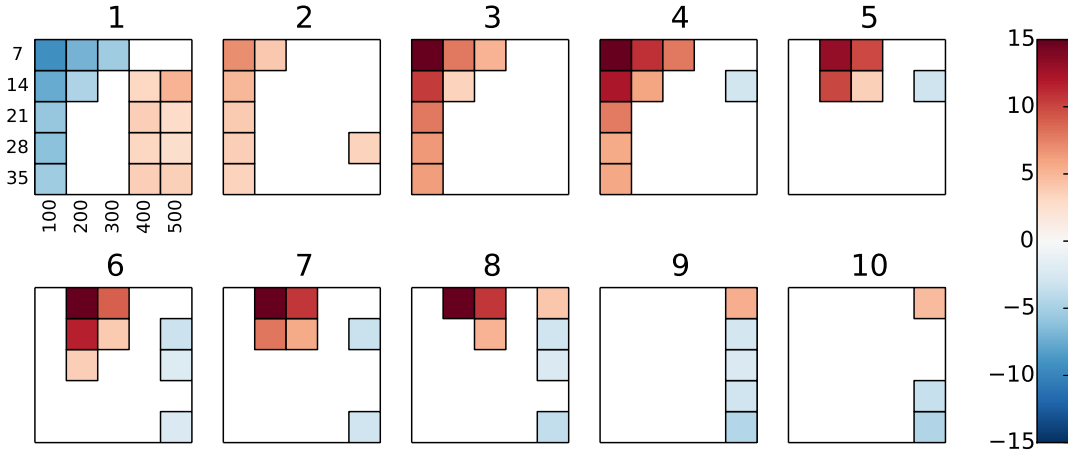
lated close pairs for which the sequence cannot be extended. This suggests that, at these scales, there is a tendency for sequences of incidents not to continue beyond the first pair and develop into more sustained bursts of offending.

Upon initial consideration, this result may appear to conflict with those found for motifs; specifically, the identification of highly-connected motifs. On closer inspection, however, this is not the case. For example, the most significant motif - motif 4 in the 3-vertex case - does itself contain a chain of length 1 (as well as a chain of length 2), and so its significance is not contradictory. If anything can be said, it is that the motifs do not tend to occur as part of longer chains (though to rigorously establish this would require further investigation).

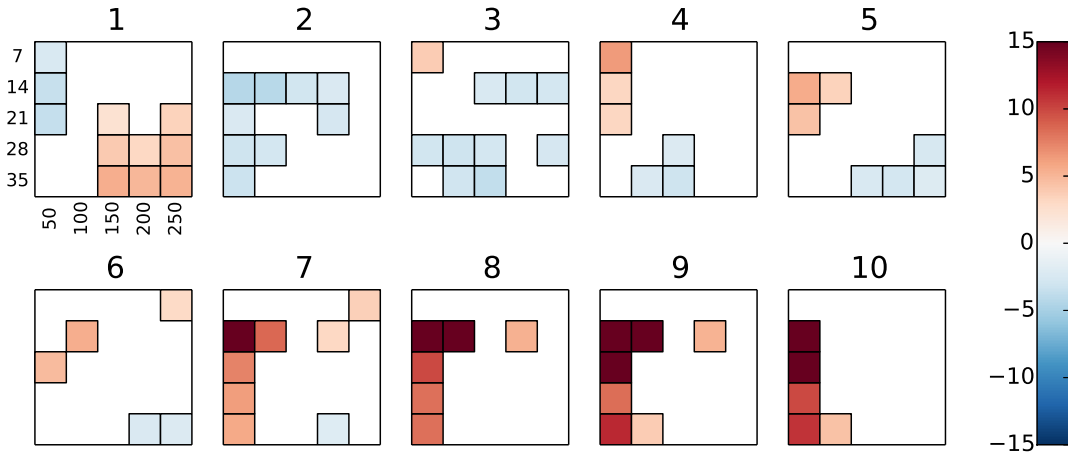
Considering Figure 2.12 again, longer length chains are found to be under-represented for both crimes. The implication of this is that, at these scales, more chains of these lengths would be expected by random chance, and that there is therefore a particular tendency for chains of such lengths not to occur in the observed data. This implies one of two things: observed chains must have a tendency to be either longer or shorter than these under-represented lengths. Given the over-representation of chains of length 1 (and the near-significant result for length 2 in the burglary case), the latter explanation is more likely.

As with the motif analysis, results for all thresholds can be summarised in a single figure, and this is done in Figure 2.13. As was the case for motifs, this reveals differences in behaviour between the two crime types. For burglary, chains of length 2, 3 and 4 are observed at relatively short spatial and temporal scales, implying that these might be typical lifetimes for spates of near-repeat victimisation at a very local level. Longer chains, of up to 8 events, are then observed at higher spatial radii (200-300m), implying that somewhat larger localities can be subjected to more sustained series of offending. There also appears to be a small effect where longer chains are inhibited at larger spatial radii, which may imply that areas beyond a

certain size are not subjected to bursts of this sort.



(a) Residential burglary



(b) Maritime piracy

Figure 2.13: Statistical analysis of event chains for a variety of spatio-temporal thresholds. For any given cell, a black outline indicates that the frequency of chains of the length is significant at 0.01 level; in such cases, the colour of the cell corresponds to the z-score for the effect, with blue indicating that such chains are under-represented and red that they are over-represented.

Results for piracy differ somewhat from the burglary case. Shorter chains, of 2 or 3 links, are under-represented for a large proportion of scales, implying that chains which reach this length tend to be sustained for a longer period. A noticeable block of significant results, however, is present for higher chain lengths at a spatial radius of 50km, across several temporal thresholds. This indicates the presence of relatively long strings of offences within the data, in which consecutive events are contained within a narrow spatial range. One way in which this can be reconciled with theory

is by reference to the use of ‘motherships’ within piracy (see Marchione & Johnson, 2013). These are larger vessels which maintain a fixed position for some period and which act as a base for attacks in their vicinity using smaller craft. The patterns arising from this, which would involve a burst of tightly-clustered attacks, all due to the same perpetrator, would be consistent with the presence of chains such as these.

2.5 Discussion

The question of whether a set of events, embedded in time and space, is clustered is a fundamental one in the study of crime, and also arises in a number of other disciplines. The statistical techniques employed when investigating the issue are typically based on pair-wise comparison of events, and identify situations in which the near co-occurrence of events in space and time is more (or less) frequent than would be expected on the basis of chance. Though these are certainly sufficient to identify clustering *per se*, their treatment of pair-wise relationships as independent occurrences means that they are unable to discern more subtle underlying patterns in the data.

In this work, the general principles of such methods are extended in order to examine the relationships between groups of events, rather than just pairs, in a way which is inspired by techniques from network theory. Specifically, the analysis of network motifs, which seeks to identify the fundamental ‘building blocks’ of networks, is translated to this context, with the corresponding aim of identifying fundamental signatures of events in space and time. In a similar vein, the analysis of chains also provides a means of examining the typical lifetime of event clusters.

The concept of an *event network* provides the bridge between event patterns and network theory. Expression of close pair relationships in the form of a graph allows all possible combinations to be examined, with the temporal precedence of events also taken into account. Existing approaches can be formulated in these terms (so that event network analysis represents a refinement of them), but the examination

of more complex features affords significantly more detailed understanding of the patterns present.

The fact that event networks are a distinctive sub-class of networks, however, presents several technical challenges. Specifically, the constraints imposed by the spatial and temporal nature of the data mean that the generation of randomised networks, which are required for statistical comparison, cannot be performed using established techniques. In addition, the fact that the analysis is intended to examine clustering *over and above* that which can be determined on a pair-wise basis means that networks generated for comparison must have prescribed properties.

Such difficulties can be overcome using an algorithm for random network generation which, in turn, is based upon the random generation of events themselves. Using an iterative scheme, such events are manipulated until the corresponding randomised network has the required properties that it can act as a suitable null comparison for the true network. In this way, the validity of the network as an encoding of feasible spatio-temporal data is guaranteed by construction, as is the level of pair-wise clustering present in the randomised data. The features of the network can therefore be analysed by comparison with this baseline.

2.5.1 Crime patterns

Application of these techniques to data for two contrasting crime types - residential burglary and maritime piracy - reveals that a number of network features occur in observed data with significantly greater frequency than would be expected by chance. This is the case for both motifs and chains, which correspond to different aspects of the clustering patterns and can be regarded as complementary analyses.

When 3-vertex motifs are considered, significant over-representation of the most highly-connected 3-vertex subgraph is seen for both datasets, from which it can be inferred that not only do events tend to co-occur, but that close pairs of events

themselves tend to group together in a highly dense manner. The fact that the over-representation of this effect is far greater than that of other possible subgraphs suggests that the majority of clustering is of this particularly tight form, in comparison with others, such as drift or convergence.

Significant motifs are also seen in the 4-vertex case, again reflecting high levels of clustering. Intriguingly, however, and in contrast to the 3-vertex case, results differ markedly between the two crime types, in terms of the particular motifs identified. These differences are consistent with the presence of distinct targeting mechanisms, or the hypothesis that different numbers of actors are involved. This is a crucial outcome from the perspective of the method itself, since it suggests that it can be used to identify differences in the patterns; traditional analysis would simply conclude that both are clustered. Having controlled for purely pair-wise clustering, the method used here reveals that the agglomeration of pairs into larger clusters can take different forms.

Translating into more explicitly spatio-temporal terms, the results imply that events tend to occur in the overlapping vicinity of previous events. This has immediate practical implications: this overlapping region, which will typically be substantially smaller than the circular region which would be identified for a single event, experiences greater elevation of risk than that associated with a single event. In a sense, risk acts additively in this region, and so a restricted area of ‘super-risk’ is present; this would be an appealing, and well-defined, target for policing activity. In addition, this result is encouraging from the perspective of crime modelling: diffusion-based models of crime, which would imply extra risk in the common vicinity of two recent events, are consistent with this finding.

When considering the particular motifs identified, descriptions of the patterns represented can become very intricate, especially in the 4-vertex case. Although it would be counter-productive to explore individual motifs in great detail, they can,

nevertheless, be characterised in a qualitative sense. For the two crimes considered here, those motifs which are unique to residential burglary tend to correspond to more spatially diverse patterns, which might be expected from theory. Though certainly the implications are requiring of deeper consideration, they are likely to have relevance to theories of spatial anchoring in offending and the ‘same offender’ hypothesis for repeat offending.

The analysis of chains again shows both agreement and contrast across the two crime types. In both cases, chains of length 1 are disproportionately over-represented. On one hand, this implies that isolated pairs do account for a significant proportion of more general clustering. From an alternative perspective, though, such results might also be regarded as a natural side-effect of more dense clustering elsewhere in the set of events: if close pairs bunch together in some locations, others are more likely to be relatively isolated.

Interesting results are also found for longer chains. Over-represented chains are, in general, shorter for burglary than for piracy, suggesting that there is a disparity in the typical length of bursts of activity in each case. Such results are of practical value since they may be used to estimate the lifetime of hot-spots and the number of incidents which they are likely to sustain. The results can also be reconciled with offenders’ modes of behaviour in both cases. For burglary, chain lengths could be interpreted as indicating the number of offences which can be committed in one location before perceived risk becomes too great, whereas, for piracy, they may correspond to the number of offences which might be attributed to the existence of a mothership. A possible direction for future work would be to examine these features in detail to examine whether these hypotheses are supported by qualitative evidence. In addition, the application of these techniques to data for other crimes which are known to exhibit clustering may also be productive.

2.5.2 Application in other domains

Although the techniques presented here were motivated by questions relating to crime, the methods themselves apply to any generic set of events in time and space. Because of this versatility, it is likely that such techniques could be applied productively in other fields for which issues of clustering are of theoretical or practical importance.

A natural candidate for this is epidemiology, in which space-time clustering can be taken as evidence of an infectious process. The features examined here have clear analogies in that domain: motifs correspond to the density and directionality of transmissions, whereas chain lengths might indicate the typical lifetime of outbreaks. Again, an important point to note is that pair-wise clustering is controlled for in identifying these features; that is, they provide more refined characterisation of patterns which are already known to arise from contagion. The fact that the techniques are capable of discriminating between clustering patterns is also relevant in this case, since it suggests that it may be possible to identify distinct modes of transmission.

Aside from those in epidemiology, there are a number of other fields in which the same principles may apply. Point process models, which can give rise to spatial contagion, have been used in the context of both seismology, in the study of after-shocks (Ogata, 1998), and ecology, as applied to species invasion (Neyman, 1939). The latter is a particularly appealing candidate for analysis since it is not a purely physical process - it is also determined by behavioural factors - and so may give rise to variability which can be well-characterised by these methods.

2.5.3 Extensions to the method

There are a number of ways in which the techniques introduced here could be refined in future work, with respect to the details of the technique itself and its potential uses. On one hand, there is clear scope for the investigation of network features

other than the motifs and chains considered here. The key theoretical innovation - the production of randomised event sets - provides a baseline against which any network property can be compared; there are likely to be several of these which correspond directly to spatio-temporal phenomena.

In addition to this, it may also be productive to explore the use of weighted event networks. The networks considered here are derived from the binary ‘close pair’ relationship and are therefore unweighted; while this is parsimonious and conceptually simple, it does imply a substantial loss of information. This could be remedied by the use of weighted links, which would represent proximity in a more granular form, and for which analytical methods do exist. The interpretation of network features in terms of real-world patterns would, however, be significantly more challenging. An alternative approach would simply be to introduce a number of spatial bands, as is done in more advanced forms of the Knox test. In this case, several event networks would be produced, corresponding to different levels of separation.

Finally, there is also an opportunity to adapt the approach from what is purely a means of statistical measurement to a predictive method. Pair-wise clustering analysis has inspired several methods of crime prediction based on the identification of a radial region around an initial event, in which a close pair might be expected to occur. 3-vertex motifs, however, correspond to a situation where a close pair of events have occurred and a third can be predicted to happen in some vicinity common to both (which, crucially, may have much smaller area than the corresponding radial region). Predicting on the basis of groups of events in this way might therefore have great potential as a predictive tool, and will be the subject of additional investigation.

Chapter 3

Quantifying the relationship between street network structure and burglary risk

The techniques presented in Chapter 2 were concerned with the characterisation of patterns in sets of spatio-temporal crime data, and the use of such techniques to draw inferences regarding criminal behaviour. Although particular crime datasets were considered in the empirical work, however, the technique itself is a general one, applicable to any set of events occurring at known locations in space and time. When using it to examine data, no use is made of contextual information and the analysis remains agnostic to the particular spatial setting. Both theory and empirical evidence, however, suggest that the spatial distribution of crime is influenced to a significant extent by environmental factors (see Brantingham & Brantingham, 1981; Andresen, 2014). The aim of this chapter is to examine the influence of one such factor: the structure of urban form, and that of the street network in particular. Previous research (*e.g.* Beavon *et al.*, 1994; Johnson & Bowers, 2010) has suggested a relationship between street network properties and the distribution of crime; however, many of the network measurements used in those studies do not correspond well to the underlying criminological theory. This chapter examines the use of more sophisticated network properties in accounting for the heterogeneous distribution of crime at street segment level. These properties are used to measure concepts such as usage potential and permeability in an objective and quantitative way, and thereby provide more well-grounded support for related hypotheses.

3.1 Introduction

Several prevailing theories of crime stress the importance of considering contextual factors when analysing the spatial distribution of offences. The general approach of ‘environmental criminology’ (Brantingham & Brantingham, 1981), for example, is based on the notion that crime can be best understood by considering the criminal act itself, and the decision process of which it is the result. Such decision processes take into account the characteristics of potential targets and are also influenced by surroundings, in terms of both the built environment and the activity occurring therein. At an even simpler level, though, for such a decision process to even arise requires that the opportunity in question has become known to the potential offender; he or she has, in some sense, encountered it.

Such reasoning implies that understanding the human activity patterns which occur in urban environments is essential to any analysis of the spatio-temporal distribution of crime (Felson, 1987). These affect both the extent to which certain opportunities are encountered and the social forces present when they are (which may affect whether the opportunity is taken). Imbalances in either of these ought to imply heterogeneity in levels of crime occurrence, all else being equal.

In the case of crimes which take place in urban areas - as will be considered here - the routine activities in which people engage, and during which opportunities are encountered, are predominantly those which include movements from place to place within a built environment, often between those which serve particular functions. The way in which the urban space is configured, therefore, shapes these journeys and determines the extent to which inhabitants become aware of the spaces around them (see Brantingham & Brantingham, 1993b). According to this argument, it should be possible to reconcile the criminal character of places with their situation in the urban form.

When considering urban form and its relationship with crime, a natural object of

study is the street network. The street network acts as a ‘skeleton’ for an urban area, in the sense that it is the structure around which all elements of the built environment are arranged. In addition, it acts as a constraint on travel and determines the paths which can be taken in moving between any two locations. On that basis, variations in the level of activity within an urban area can be assessed most appropriately in terms of street network usage.

For crime in particular, it is also the case that many criminal events occur at some location on the street network: as well as being the substrate for movement, it is also the structure on which targets are arranged. This is particularly clear - and analytically expedient - in the case of crimes for which the target is located at a fixed point on the network. Burglary, which will be considered here, is such an example. More generally, analysis based on the street network is well-aligned with the use of small spatial units in geographic criminology, as advocated by several scholars in the field (see, for example, Weisburd *et al.*, 2009); this assumes, of course, that a sufficiently granular network representation is used.

Analysis concerning the street network is also advantageous, and of significant potential value, from a practical perspective. The units at which crime reduction activities are implemented are commonly those which can be specified in terms of the street network, *e.g.* street segments or named roads. Police patrols, which are specified in terms of routes and would themselves involve travel around the network, are an example of this. This is also the case for potential uses of such analysis not directly related to crime reduction: insurance premiums, for example, typically take neighbourhood characteristics into account, often specified at street level.

As would be expected on the basis of these observations, the relationship between crime and the street network has been the subject of several empirical investigations (*e.g.* Beavon *et al.*, 1994; Hillier, 2004; Johnson & Bowers, 2010). Though these differ in their precise emphases and findings, they demonstrate consistently

the existence of significant relationships between network structure and crime phenomena. In most cases, these are then reconciled with theories of offender awareness and decision-making, as well as the role of guardianship.

Nevertheless, these approaches suffer from certain shortcomings related to their measurement of the street network. Administrative road classification, for example, is used in many studies, but is not necessarily consistent and, as a categorical variable, suffers from a lack of granularity. Although more explicitly quantitative variables have also been considered, many of those used, such as connection counts, have an association with activity patterns which is somewhat opaque. This is particularly problematic when such findings are used to make inferences about the behaviour of offenders.

The quantitative analysis of street networks, however, has become significantly more advanced in recent years (see Crucitti *et al.*, 2006b), in line with the growth of network science more generally. In particular, this has involved the introduction of several metrics which are particularly appropriate to street networks (which have distinctive character, in comparison with networks in general). Moreover, several of these metrics have an immediate interpretation in terms of travel patterns on such a network, relating in particular to the accessibility and likely use of individual places.

Since these concepts are precisely those which are of interest from a criminological perspective, such metrics provide an opportunity for analysis of crime which is both quantitatively sophisticated and can be related directly to criminological theory. The aim of this chapter is thus to extend understanding of street network effects on crime by employing an explicitly quantitative analytical framework for the study of urban streets, thereby formalising several of the general ideas which have been explored previously. This chapter presents such analysis, considering in particular the relationship between the network metric ‘betweenness’ and levels of burglary. In the process of this, several issues related to street network analysis in

this context will also be discussed, and empirical results concerning network clustering more generally will also be presented.

Before presenting the empirical research, a variety of background material will be reviewed. In the first case this will concern the theoretical background from a criminological perspective, before moving on to introduce the technical study of street networks and previous work in the field.

3.1.1 Crime and urban form

Routine activity theory (Cohen & Felson, 1979) considers the ecological conditions that give rise to crime events and has been highly influential in theories of crime pattern formation. In its simplest form, the theory states that direct-contact crimes can only occur when a motivated offender encounters a suitable target in the absence of a capable guardian (that would otherwise prevent the crime). Crimes take place at particular locations at specific times, and so these three elements (offender, target and absence of guardian) must simultaneously converge in order for a crime to occur. The street network is a fundamental determinant of this convergence. In the case of burglary, targets are positioned at fixed locations on the network, and their density varies from street to street. Potential offenders encounter opportunities while moving around the network and must travel to and from any offences; in both of these, the routes they take are constrained by the configuration of the network. Likewise, their non-criminal pedestrian activity (which may affect guardianship), and that of formal guardians such as the police, is not random and is shaped by their journeys through the network. As suggested by Brantingham & Brantingham (1993b), the network therefore captures many of those aspects of the ‘urban backcloth’ relevant to the location of criminal activity.

The ways in which offender awareness and guardianship are shaped by the configuration of the street network can be understood by appealing to crime pattern theory (Brantingham & Brantingham, 1993a). In the context of burglary, crime pattern

theory asserts that offenders typically choose to victimise properties encountered during the course of non-criminal activities. More specifically, offenders are said to form awareness spaces, conceived as ‘cognitive maps’, during their routine activities, and it is where these activity spaces intersect with suitable opportunities for crime that individuals are most likely to offend. Considered in these terms, the question of risk to a given property can therefore be translated to one concerning the extent to which it features in the awareness spaces of potential offenders. If it can be assumed that more awareness corresponds to higher risk, then understanding the aggregation of awareness spaces becomes crucial to the study of the distribution of risk.

Since awareness spaces are formed during the course of regular activities, their location and extent are determined by the travel patterns arising from such activities. Crime pattern theory emphasises the role of individuals’ primary ‘activity nodes’ (*e.g.* homes, workplaces and leisure premises) in shaping these, with awareness spaces thought to be centred around these or along the journeys between them. A straightforward corollary of this is that places which are activity nodes for many people are therefore the intersections for many individual awareness spaces and ought, accordingly, to have characteristic crime patterns. Empirical research has supported this, showing that such locations (*e.g.* transportation hubs and entertainment districts) experience elevated levels of various crime types, including robbery (Bernasco & Block, 2011), drug crime (McCord & Ratcliffe, 2007) and alcohol-related crime (Block & Block, 1995). By similar reasoning, the fact that certain roads are likely to be used more than others during routine trips (the hierarchical nature of street networks means that certain roads act as conduits for a large proportion of journeys) suggests that they ought to feature in the awareness spaces of many people, including potential offenders. Since these spaces determine the criminal opportunities of which they become aware, it is to be expected that the extent to which they overlap corresponds to the spatial distribution of potential for crime.

Turning to the issue of guardianship, the theoretical context is significantly more

complicated, and the subject of some debate (Mawby, 1977; Hillier & Shu, 2000; Reynald & Elffers, 2009). Pedestrian activity is, again, a crucial issue, since the number of potential guardians at a given time and place will be determined to a large extent by levels of throughput. However, although the reasoning discussed above - that patterns can be understood in terms of journeys through urban space - can again be applied to understand variation in this, the effect on crime is not necessarily clear.

On one hand, it is suggested that increased pedestrian activity has the effect of increasing vigilance and therefore acting to reduce crime. This effect, succinctly encapsulated by Jacobs (1961) as the notion of “eyes on the street”, may counteract the effect associated with increased offender awareness of busy places by providing a ready supply of guardians, the prospect of whose intervention dissuades potential offenders from committing crimes. Were this to be the case, less crime would be expected on those streets with the highest usage; this is a testable hypothesis in network terms. In addition, the implications of this from the perspective of planning are clear: that urban design should encourage pedestrian throughput as an indirect and natural method of crime control.

Some authors have, however, suggested that this effect is overstated and insufficient to offset the increase in risk associated with greater exposure. The willingness and ability of passers-by to intervene has been shown to be a complex issue (Reynald, 2010), and may not be as significant an impediment to criminal acts as might be expected. In fact, an alternative school of thought suggests that the converse is true: that transient pedestrian activity (in sufficient volume) simply serves to diminish the territoriality of places and therefore renders them less daunting to potential offenders, particularly those who are outsiders (Newman, 1972). Ethnographic studies with burglars (Cromwell *et al.*, 1991; Ashton *et al.*, 1998) suggest that exploratory search strategies are relatively rare, and that offenders are generally unwilling to venture to areas where they will ‘stand out’. Consequently, locations that are already

known to them and for which the ambient population is transient may be particularly attractive. In line with this, Brantingham & Brantingham (1975) looked at the configuration of neighbourhoods, finding that crime tended to be higher at boundaries, where community structure is likely to be less well-defined. This finding is also consistent with theories of collective efficacy (*e.g.* Sampson *et al.*, 1997), which argue that stronger, more cohesive communities - where residents know each other and have a shared interest in the space - are better equipped to defend against and deter intruders.

Several empirical studies have examined various aspects of the relationship between the street network and burglary. The earliest is the work of Bevis & Nutter (1977), who examined burglary levels on isolated streets and studied the effect of network density at the census tract level, using data from Minneapolis, USA. The concept of density is defined in that work as the ratio of the number of street segments to the number of junctions. This is a relatively coarse-grained metric, but nevertheless a relevant one: a minimal functioning street network would be tree-like in structure, with relatively long trips required to reach most destinations, and would yield a low score on this metric. Increasing density via the addition of more street segments adds redundancy and therefore plurality in terms of travel path choices, with the distribution of throughput becoming more balanced as a result. A regression analysis presented in the study revealed a positive relationship between area-level density, defined in this way, and burglary levels, after controlling for various social and demographic variables. Individual streets which were categorised as inaccessible on the basis of the number of connecting segments were also found to experience less victimisation, on average, than their more connected counterparts.

Also working at the area level, White (1990) examined burglary risk across 86 neighbourhoods in Massachusetts, USA. In this case, the independent variable was the number of roads in each neighbourhood that were connected to (categorically-defined) major roads. While offering relatively little insight into traffic flow, such

a metric corresponds most obviously to permeability in the sense that it provides an indication of the ease of access of neighbourhoods. As with the Bevis & Nutter study, it was found, after controlling for various other relevant factors, that network structure was a statistically significant factor in determining area-level crime rates, with more accessible areas having higher burglary rates.

At a lower spatial scale - the street segment - Beavon *et al.* (1994) considered levels of crime risk (including, but not limited to, burglary) in Ridge Meadows, Canada. They used two metrics to capture different network effects: for a given segment, accessibility was measured as the number of other segments with which it shared a junction, whereas flow was estimated via a hierarchical road classification system. After again controlling for various other factors in statistical analysis, both factors were found to be positively associated with crime: risk increased with the number of connected streets, and roads classified as being more important also experienced greater victimisation.

The role of the street network in urban processes is one of the primary foundations of ‘space syntax’, an approach that has been applied in a variety of urban contexts (Hillier, 1996). Much of this work rests on the fundamental observation that individual street segments cannot properly be understood in isolation, but must be considered in the context of the entire network and their position within it. In addition, it places particular emphasis on lines of sight as a determinant of connectivity (and, in some cases, as the definition of the street segment itself), which, as noted previously, is of particular relevance in the context of crime. Several metrics have been developed using these ideas, such as ‘integration’, which measures how close a given location is to all others in terms of paths through the network. Applying these methods to crime, and analysing data from London, UK, and an urban area in Australia, it was found that crime was positively related to connectivity but negatively related to integration (Hillier, 2004). The conclusions drawn from this are that permeable designs are favourable, but that the effect can be reversed

where redundant connectivity is present (that which does not increase integration), perhaps because of the provision of extra entry or escape routes.

The space syntax work also focuses on one particular class of street segment: the *cul-de-sac*. Such segments represent a neat example for analysis since they can be universally defined, relatively independently of spatial context, and represent an extreme case in theoretical terms, since they will usually be remote and experience little through traffic. In the analysed data, risk on these segments was found to be considerably higher than on others, although it is emphasised that physical shape must also be taken into account. Linear and well-connected segments were found to be safe, whereas those which were sinuous and/or relatively secluded were at higher risk (Hillier, 2004). Clearly this is a finding that runs contrary to what would be anticipated according to pattern theory.

Hillier's work makes several other observations at the level of individual properties, in particular relating to their modes of access (*e.g.* the direction they face or proximity to alleys). Such features were also considered by Armitage (2007), whose work involved detailed assessment of the physical features of individual houses. These observations included the type of road on which they were situated and a subjective estimate of its usage, both of which were then compared individually with crime levels. Once again it was found that increased activity and permeability, as measured in these senses, were associated with higher risk. One particularly clear example of this arose from comparison between isolated *cul-de-sacs* and those which were serviced by 'leaky' footpaths (which ought to increase permeability): risk was found to be greater for the latter, and this provides further support for the overall finding.

A subsequent paper by Johnson & Bowers (2010) explored many of these issues using a statistical framework specifically designed to account for the hierarchical structure of the data; that is, the nesting of street segments within neighbourhoods, and that of neighbourhoods within larger areas. In addition, their statistical ap-

proach differed from previous studies by accounting for the fact that, while crimes are rare, incident counts may exceed 1. Using burglary data from Merseyside, UK, they considered variables similar to those seen in the previous work at the street segment level - number of connections, road classification and physical shape - with a particular emphasis on *cul-de-sacs*. Their hierarchical linear analysis showed that, in addition to significant variation at higher levels of spatial aggregation, there was a positive effect of connectivity on crime, and that higher rates of victimisation were observed on major roads (all else being equal). *Cul-de-sacs* were found to be the safest type of street segment, although a marked difference was observed between those which were sinuous in form (which had lower crime rates) and those that were linear. The findings of this study support the general hypothesis that permeability is associated with higher crime, but studies of this scale and that use appropriate statistical methods are rare.

Though the distribution of crime is not its primary focus, the work of Iwanski *et al.* (2012) is also notable for its treatment of the street network. Working with data for offenders' homes and offence locations, the work takes as its basis the hypothesis that the crimes studied were committed while the offender was in the process of a longer journey. An estimate is produced for the onward path he or she may have been taking, where navigational choices are based on a combination of directionality and the likely popularity of roads (which is estimated by simulating a number of journeys through the network). Results show that these paths tend to lead towards known 'crime attractors': locations towards which offenders are drawn due to a known abundance of criminal opportunities, such as shopping centres and transit stations (Brantingham & Brantingham, 1995). It is worth noting, though, that the findings are also consistent with a somewhat simpler hypothesis that crime attractors tend to be located on roads with high activity.

3.1.2 Shortcomings of existing approaches

The analytical work described above demonstrates consistently that significant relationships exist between the street network and crime patterns, and thereby confirms that the distribution of crime can be meaningfully understood in those terms. However, a number of limitations can be identified among the methods used, which suggests a need to consider the exact nature of the relationships identified. In particular, these concern the way in which street network properties are measured, in terms of the level of detail which they afford and their appropriateness as measures of the underlying quantity of interest.

Several studies, for example, use categorical road classification (*e.g.* major/minor, ‘A/B Road’, and so on) as a proxy measure of the activity level on a given street (*e.g.* White, 1990; Beavon *et al.*, 1994; Johnson & Bowers, 2010). This is problematic in several respects. Firstly, it is an ordinal variable with a small number of possible values: few classification systems have more than 10 meaningful categories, and this can be reduced further in cases where certain types can be discounted as potential crime locations (such as motorways, in the case of burglary). Considerable variation is likely to be disregarded when such a coarse classification is used.

This problem is exacerbated further by the fact that such categories do not correspond consistently to the kind of activity levels they are intended to measure. While certainly broadly indicative of likely usage, road classifications are primarily an administrative construct and numerous other factors are taken into account in their specification. In addition, they are defined for roads in their entirety and do not capture variations between different sections of the same road, which can be considerable; to assume that they do is a form of the ‘ecological fallacy’ (Robinson, 1950). Real-world examples are readily available of pairs of roads for which the relative levels of usage conflict with those which would be expected on the basis of classification (see also Section 3.2.2.3).

One further point concerns the wider applicability of such research, particularly as regards inter-national comparison. Road classification systems differ from country to country, often with no natural correspondence between categories. Indeed, certain classes are specific to a limited number of countries (toll roads, for example) and complicate the situation still further. Comparison between nations on this basis is therefore futile in all but the most general terms.

Of course, a number of objective and highly spatially-granular variables, for which these concerns do not apply, are also considered in the aforementioned studies. For several of these, though, their interpretation in terms of human activity is open to question. One example is the counting of adjacent roads; that is, for a given street segment, counting the number of other segments with which it shares a junction. Although this is indicative of the *potential* for access, in some sense, it reveals nothing about the probability of such access. It is also a local measure, and does not take the wider spatial context into account, which is at odds with the principles of space syntax (Hillier, 1996), for example. A street may have a large number of connections, but, if its location is peripheral in terms of the wider network, it is unlikely to be highly-used.

Similar issues can be identified with a number of other variables used in previous analyses. In order to address these problems, it is clear that it is necessary to explore the use of alternative variables which are better suited to the present purpose. As a wider point, the discussion also highlights the need to take particular care when selecting variables, not only in terms of their suitability for meaningful statistical analysis, but also as regards their relationship with theoretical concepts. In the case of crime analysis, the mechanism by which the street network exerts an influence is clearly specified - as shaping the accumulation of awareness spaces during routine journeys, for example - and it is crucial that network measures relate clearly to this.

3.1.3 Street network analysis

Recent years have seen a surge of interest in the study of networks, in particular as they relate to the structure and dynamics of complex systems (an introduction is provided by Newman, 2010, and basic concepts were outlined in Section 1.4). Throughout the diverse range of applications of network science (see Boccaletti *et al.*, 2006), one of the primary themes is the empirical measurement of networks using an array of metrics designed to address various aspects of their structure. Though many of these have been appropriated from other fields, such as social network analysis (Wasserman & Faust, 1994), others have been proposed with novel applications in mind, such as *Google's* ‘PageRank’ (Page *et al.*, 1999).

Street networks have been the subject of research throughout the development of network theory. Indeed, one of the classic problems of graph theory, the ‘Königsberg Bridge Problem’ of Euler (1741), is inspired by a question related to road layout. For analytical purposes, their study lies within the sub-field of ‘spatial networks’, which is a distinctive genre requiring bespoke treatment. In contrast to more abstract networks, the fact that vertices are embedded in space imposes certain constraints on the possible structures which can arise, thereby complicating analysis. In particular, the fact that there is typically a cost associated with longer-distance links biases their formation in favour of shorter links and tends to give rise to more ‘regular’ structures.

In addition, physical constraints mean that some structures are simply unrealistic for spatial networks. To take the example of rail networks, there is a physical limit to the number of tracks which can meet at a station: in network terms, this corresponds to an upper limit on the vertex degree. This is in contrast to other types of network, for which no such constraint exists: in the case of the web hyperlink network, for example, there is no such issue. For reasons such as this, some classic metrics, such as degree, are relatively uninformative for spatial networks, due to the narrowness of their distributions. Alternative metrics, typically based on notions of

travel through the network, are therefore more common in such cases. A full review of the field of spatial networks is provided by Barthélemy (2011).

3.1.4 Mathematical representation

In order to understand the various metrics studied in the context of street networks, it is first necessary to introduce the methods by which such networks can be represented mathematically. The natural approach for doing this is by encoding its structure in a graph; however, there is more than one way to do this. The various representations emphasise different aspects of the network structure, and their suitability depends on the perspective of the particular research application. The two most common of these - the *primal* and *dual* representations - will be introduced and discussed here. The empirical work presented in the later part of this chapter focusses on the primal representation, and this choice is discussed below.

3.1.4.1 Primal representation

The *primal* representation (Porta *et al.*, 2006b) is the more intuitive of the two representations, and corresponds most closely to a traditional street network map. In such a graph, each junction in the street network is represented by a vertex (where ‘junction’ refers to any point at which a traveller through the network would have a choice of path). A link is then added between any pair of vertices for which the corresponding junctions are connected directly by a street; such a section of street, connecting two junctions, is defined as a *street segment*. Figure 3.1 shows the stages of construction: the identification of junctions and the addition of links between them.

The term ‘primal’ refers to the matching between dimensionalities of features: junctions (which are points, and therefore zero-dimensional) are represented by zero-dimensional vertices, and street segments (linear and one-dimensional in space) are encoded as links, which are also one-dimensional. Because of this, each graph feature can be given meaningful spatial attributes: vertices have a definite geographical

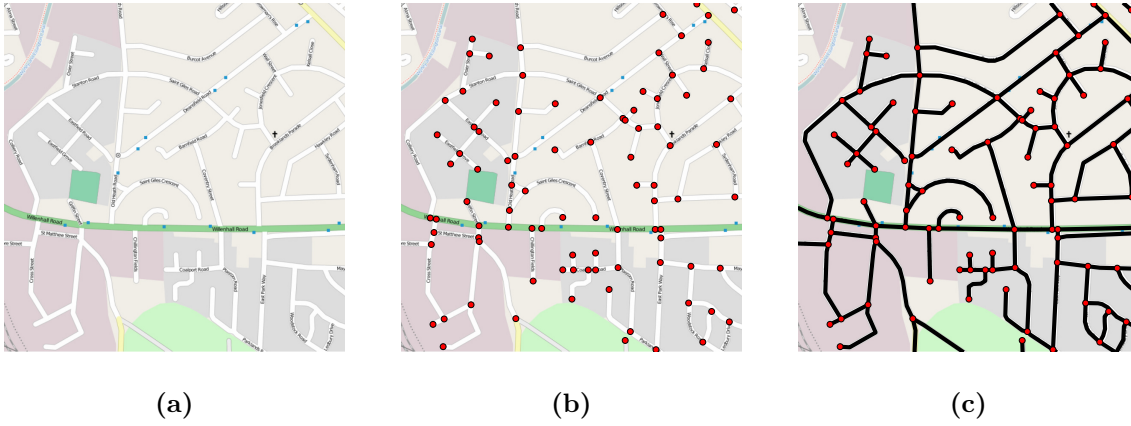


Figure 3.1: The construction of the primal representation of a street network: a) the original map; b) vertices placed at each junction; and c) links added between any pair of junctions connected by a street segment.

position, and links have a physical length. The latter of these means that this can be regarded as a *weighted* graph, and the physical length of any route through the network can be found by reference to the graph.

3.1.4.2 Dual representation

The *dual* representation is that which arises when the roles of vertices and links are inverted, relative to the primal case. Streets are represented by vertices, and two vertices are linked if the associated streets intersect at some point. That vertices represent ‘streets’ rather than ‘street segments’ is significant: in this context, a street is a section of the network (possibly comprising multiple segments) which can be regarded as a coherent single entity. The intersections in question are therefore those between these larger unified streets.

There are several options for the process by which unified streets are identified. In the ‘space syntax’ approach (Hillier & Hanson, 1984), which pioneered representations of this type, streets are defined on the basis of ‘axial lines’: straight lines which can be interpreted as lines-of-sight. As noted by Porta *et al.* (2006b), this is only obliquely related to the street network configuration itself, since the intersections of axial lines do not necessarily coincide with junctions. A further problem arises because such representations may not be unique: for a given street network,

valid axial lines can be drawn in multiple ways.

An alternative is the ‘named street’ approach (Jiang & Claramunt, 2004), whereby contiguous elements of the network are unified if they share a common street name. A clear shortcoming of this nominalistic approach is that it depends on the reliability (and availability) of street names, which can be subjective and arbitrary. As an alternative, Porta *et al.* (2006a) proposed an algorithm in which street segments are grouped according to their geometric linearity at junctions, so that a street is a relatively continuous sequence of segments.

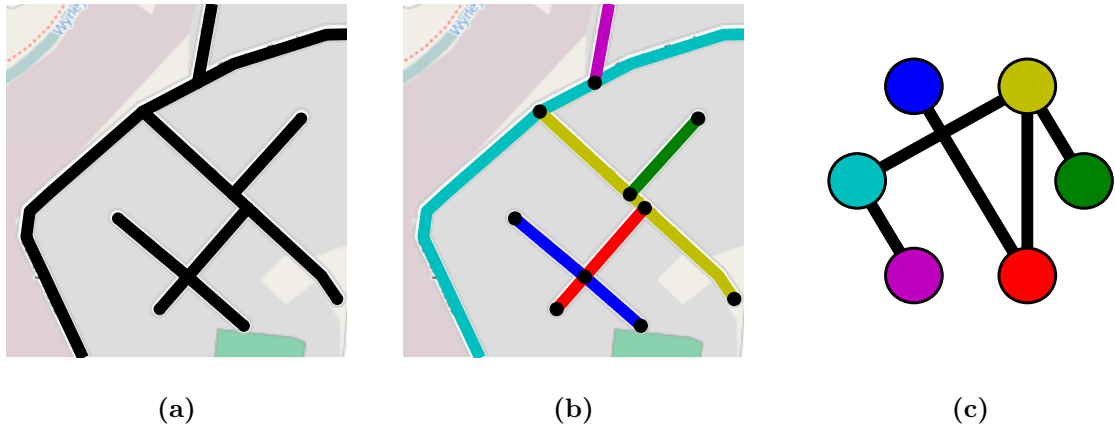


Figure 3.2: The construction of the dual representation of a street network: a) the original map, with black lines placed along streets; b) streets identified and coloured on the basis of street name; and c) the dual network derived on the basis of street intersections. The latter network is aspatial: other than indicating adjacency, the location and form of vertices and links has no geographical meaning.

Figure 3.2 shows the construction of a dual graph in which, for ease of presentation, the named street approach is taken. Segments are coloured according to street name and those which intersect are linked in the derived graph. Its aspatial nature - a result of the dimensional mis-match - is immediately apparent: streets are collapsed to vertices, for example, which have no property analogous to length.

3.1.4.3 Choice of representation

Arguments in favour of the dual approach typically focus on its relevance to issues of wayfinding in an urban environment. Rosvall *et al.* (2005) express their relevance in informational terms, as corresponding to the way in which travellers perceive and encode journeys through the network. The information required to navigate from one point to another essentially comprises a list of important ‘turns’ (changes from one street to another), rather than a complete enumeration of all street segments. The number of such turns corresponds precisely to the notion of path length in the dual network (which is a topological, rather than metric, measure): each link traversed corresponds to a turn. When streets are based on some form of linearity, this is also well-aligned with the observed preference of wayfinders to go straight at intersections (Conroy Dalton, 2003).

There are, however, several shortcomings of the dual approach. Although concerned specifically with inconsistencies in space syntax, Ratti (2004) makes several points related to dual representations more generally, such as their susceptibility to edge effects and the loss of important geographical information. This is reaffirmed by Porta *et al.* (2006b), who suggest that the loss of metric information is simply too big a price to pay in the context of geographical analysis.

The issue of street unification is also problematic when the ultimate aim of the analysis concerns other processes, such as crime. A single street (represented by a vertex) in the dual representation can be very long and comprise many street segments. Treating such a large feature as a single geographical entity - which would imply distilling all crime to a single value - conflicts directly with the aim of carrying out granular analysis.

Although problems can also be found with the primal approach - it is at risk of placing undue emphasis on the role of junctions, for example - the arguments above have led to it being favoured in recent studies (Crucitti *et al.*, 2006b; Porta *et al.*,

2009; Chan *et al.*, 2011). In the remainder of the analysis presented here, the primal representation will be used, with similar justification: namely, its granularity and explicitly geographical perspective. Though the dual representation will not be considered further here, its properties have been investigated in a number of empirical studies (Jiang & Claramunt, 2004; Porta *et al.*, 2006a; Kalapala *et al.*, 2006; Jiang, 2007).

3.1.5 Physical analysis of street networks

Recent developments in the study of street networks have focussed on a number of distinct aspects, including their geometric properties, their evolution and their graph-theoretic properties. These provide valuable contextual information regarding the suitability of various measurements for application to crime analysis.

Although not the focus of the present work, the geometry of networks, and their basic graph-theoretic properties, has been the subject of a considerable volume of research. The nature of street networks means that they are a particularly interesting case: as well as being spatial (as detailed above), they are also *planar*. Planar graphs are those in which the links do not physically cross; a characteristic which has implications for their properties (although strictly the presence of bridges and tunnels causes this to be violated, it is approximately true in general).

The network properties investigated include basic link and vertex counts (and their ratio) and vertex degree distributions. On the geometrical side, measures typically concern segment length, total length, intersection density and the shape and area of ‘cells’ (the polygons enclosed by streets). Various combinations of these measures have been presented for a diverse range of cities (Buhl *et al.*, 2006; Cardillo *et al.*, 2006) - with other work concentrating specifically on German cities (Lämmer *et al.*, 2006; Chan *et al.*, 2011) and London (Masucci *et al.*, 2009) - and a number of scaling laws found in those studies and elsewhere (Barthélemy & Flammini, 2008).

Geometry is also the fundamental concern of a related area of research concerning the modelling of street network evolution. Suggested approaches model the emergence and growth of sections of street, and relate this to more general urban growth over time. Again, this is not the primary focus here, but several promising approaches have been suggested (Barthélemy & Flammini, 2008; Masucci *et al.*, 2009; Strano *et al.*, 2012).

3.1.6 Path-based metrics

Measurement of the graph-theoretical properties of street networks is challenging since, as noted above, many traditional metrics are not appropriate. Because of this, recent approaches have explored the use of other metrics that are better aligned to the main issues of interest for street networks. Many of these relate, in some sense, to journeys through the network, and three will be described here: betweenness, closeness and straightness.

To introduce definite notation, the network is represented by an undirected graph $G = (V, E)$, which is composed of a set of N vertices, $V = \{v_i\}$, and a set of M links, $E = \{e_i\}$. This is constructed according to the primal representation of the network: in terms of streets, N is the number of junctions and M the number of street segments. Without subscripts, v and e are taken to refer to generic vertices and links, respectively.

A *path* in a network is any ordered sequence of vertices such that every consecutive pair of vertices is connected by a link (*i.e.* a sequence of vertices which can be traversed by following links). The *length* of such a path can be defined in either metric or topological terms. The metric length of a path is the sum of the physical length of all constituent links, whereas the topological length is simply the number of links involved (the number of ‘hops’, which is 1 fewer than the number of vertices in the path). Although metric distance is intuitively the more meaningful form - its availability is one of the main motivations for using the primal representation -

topological distance does have relevance, and may correspond to notions of ‘mental distance’. The definitions of the majority of metrics are agnostic to this distinction.

For any pair of vertices v_i and v_j , a path between the two may or may not exist, and indeed there may be more than one. When a path does exist, a *shortest path* between v_i and v_j is one such path of minimal length (though, again, there may be more than one if there are multiple paths of equal length), and this length is denoted d_{ij} . By convention, d_{ij} is taken to be infinite when no path exists. To recapitulate in real-world terms, d_{ij} is the shortest distance (either in metric or topological terms) one would have to cover to travel between two junctions v_i and v_j through the street network. Efficient algorithms for the calculation of shortest paths can be found, along with many others related to graphs, in the book by Cormen (2009).

Betweenness The metric *betweenness* was proposed independently by Anthonisse (1971) and Freeman (1977), in the context of social networks, as a measure of the extent to which vertices occupy a position of ‘brokerage’. Rather than measuring how *near* they are to others, it measures the extent to which vertices act as an intermediary in communications *between* others. Two quantities are required for its calculation: σ_{jk} , which is the total number of shortest paths between two vertices v_j and v_k ; and $\sigma_{jk}(v_i)$, which is the number of shortest paths between v_j and v_k **which pass through** v_i . Using these, the betweenness centrality of a given vertex v_i is defined as

$$C_i^B = \sum_{\substack{v_j, v_k \in V, \\ v_j \sim v_k}} \frac{\sigma_{jk}(v_i)}{\sigma_{jk}}, \quad (3.1)$$

where \sim here represents the relation ‘there exists a path between v_j and v_k ’ (this simply restricts the sum to pairs for which $\sigma_{jk} \neq 0$). The essence of the metric is to count the number of times that v_i features on paths through the network, assuming that a path exists between every possible pair of vertices. A more intuitive understanding can be gained by considering how it can be calculated:

- 1) initialise all vertices with $C^B = 0$;

- 2) consider all pairs of vertices v_j and v_k ;
- 3) for each pair, find the shortest path(s) between them;
- 4) every time a vertex appears in one of these shortest path(s), increment its C^B by $\frac{1}{w}$, where w is the number of shortest paths between v_j and v_k (so if there is only one shortest path, add 1 to the C^B of each intermediate vertex).

In many applications, C^B is normalised by dividing through by its maximum possible value: $N(N-1)$, if the graph is fully-connected. This maximum is meaningless, however, if the graph is not connected and, since the primary interest is usually only in the relative values of vertices, no normalisation is performed here.

Since paths are composed of both vertices and links, betweenness can be defined equally well for links. If $\sigma_{jk}(e_i)$ is analogously defined as the number of shortest paths between v_j and v_k which pass through e_i , the *link betweenness* for a link e_i is defined similarly as

$$C_i^B = \sum_{\substack{v_j, v_k \in V, \\ v_j \sim v_k}} \frac{\sigma_{jk}(e_i)}{\sigma_{jk}}. \quad (3.2)$$

In the same way as for vertices, this is the frequency with which each link appears in paths through the network. Importantly, the betweenness of a link is not equal to the average betweenness of its two vertices: a rarely-used link can connect two highly-used vertices, for example. Whenever used, the distinction between link and vertex betweenness will be made clear from the context. The paper by Brandes (2008) presents algorithms for the efficient calculation of both forms, as well as other variations.

In terms of street networks, betweenness has a relatively clear interpretation as an estimate of the use of any given feature (junction or segment) by traffic passing through the network. Although the premise of single trips between all junctions is crude, it nevertheless represents a well-motivated first-order heuristic for urban movements. Of particular note is the fact that the value for every feature depends

entirely on its role in the wider network. The stylised example of Figure 3.3, showing link betweenness, is illustrative of its discriminatory value: the two links identified - one peripheral and one highly central - are those with values at each extreme.

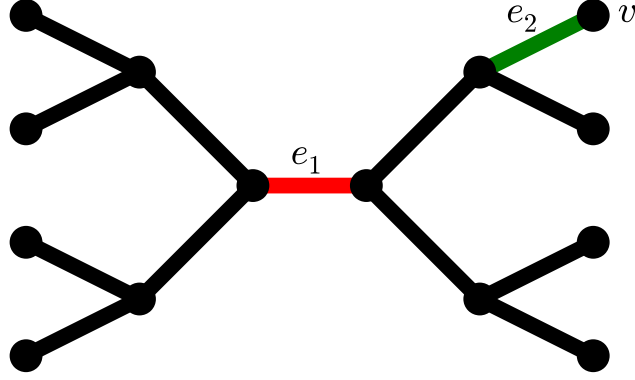


Figure 3.3: Stylised illustration of link betweenness: link e_1 (shown red) features in any path between one of the 7 vertices on the ‘left’ of the network and the 7 on the ‘right’, and therefore has a relatively high betweenness value of 49. Link e_2 (green), on the other hand, is only traversed by paths starting/ending at v ; there are 13 such paths and it therefore has a relatively low value of 13.

Closeness The *closeness* of a vertex (Sabidussi, 1966; Freeman, 1978) is a measure of the extent to which it is near to all others, on average. It is defined for a vertex v_i as

$$C_i^C = \frac{N - 1}{\sum_{\substack{v_j \in V, \\ v_i \neq v_j}} d_{ij}}, \quad (3.3)$$

so that it is simply the inverse of the average length of all paths starting at v_i . It is important to note that C^C is only meaningful for connected graphs (since otherwise $d_{ij} = \infty$ for at least one v_j), but this can generally be assumed to be the case for street networks.

In terms of street networks, closeness can be most naturally interpreted as representing the ‘accessibility’ of a vertex, in some sense. Locations with high closeness are those which are easiest to reach (in terms of distance) from all other locations, on average.

Straightness The concept of *straightness* (Vragović *et al.*, 2005; Porta *et al.*, 2006b) examines the extent to which paths in the network deviate from a perfectly linear path in space. As such, it is only valid when d_{ij} is defined metrically. When this is the case, if l_{ij} is defined as the Euclidean (as-the-crow-flies) distance between vertices v_i and v_j , the straightness of a vertex v_i is given by

$$C_i^S = \frac{1}{N-1} \sum_{\substack{v_j \in V, \\ v_i \neq v_j}} \frac{l_{ij}}{d_{ij}}. \quad (3.4)$$

When v_i is thought of as the terminus of the paths in question, this value represents the extent to which it can be reached by straight, non-convoluted paths (regardless of their length). As such, it can be regarded as representing the prominence of a location; its ‘presence’ for those travelling around the space (Conroy Dalton, 2003).

Radial limits One problem which can arise with the three metrics described above is that of ‘edge effects’, whereby misleading results can be found near the spatial extremities of the network. This is particularly acute for closeness, since vertices will tend to receive higher values the closer they are to the centre of mass of the junctions (Porta *et al.*, 2006b). To alleviate this problem, the definitions can be modified to local forms, in which only vertices within a certain distance of the vertex in question are considered in the calculation. For example, closeness is calculated as the inverse of the average distance to all vertices within some given radius, r .

In the case of betweenness, only paths of length less than or equal to r are taken into account when calculating scores (in effect, the relation \sim in equations (3.1) and (3.2) is replaced by ‘there exists a path **of length at most** r between v_j and v_k ’). These radially-limited variants are denoted by superscripts thus:

$$\left(C_i^{B(r)}, C_i^{C(r)}, C_i^{S(r)}\right) = \left(C_i^B, C_i^C, C_i^S\right) \text{ for only paths such that } d \leq r \quad (3.5)$$

The rationale for considering local measures is not purely technical. For many applications, it may be more appropriate to only consider paths of limited distance:

journeys across the breadth of a large city, for example, are scarcely relevant for pedestrian travel. The use of local measures ensures that these indices can be defined consistently and that they correspond properly to the issue of interest.

3.1.7 Research using path-based metrics

The metrics introduced above form the basis of a comprehensive approach to primal street network analysis introduced by Porta *et al.* (2006b). Their framework, known as Multiple Centrality Analysis (MCA), has three key principles: the use of a primal representation; the use of metric distance; and the treatment of centrality as a multi-faceted concept with several complementary aspects. Betweenness, closeness and straightness are three such aspects. By construction, each of these seeks to capture some different component of urban structure, and a comprehensive impression can be gained by considering them in concert.

In other studies by the same team (Crucitti *et al.*, 2006a,b), MCA was performed for 1-square-mile sections of 18 world cities, selected for their diversity (*e.g.* planned *vs.* self-organised). Within these, common functional forms were found for the distributions of the various centrality indices, as well as correlations between them. Conversely, other properties appeared to show different behaviour for different types of city. The inequalities in some of these distributions were also used to propose a typology of these cities, using a hierarchical clustering algorithm. Aside from these studies, others have also examined the particular case of betweenness centrality in some depth (Lämmer *et al.*, 2006; Scellato *et al.*, 2006).

Although the properties of these networks are interesting in themselves, of even greater interest is their relationship with activities taking place upon them. Recent research on this theme has been particularly promising, comparing the results of MCA with economic activity in both Bologna, Italy (Porta *et al.*, 2009) and Barcelona, Spain (Porta *et al.*, 2012). In those studies, Kernel Density Estimation (KDE) is used as an analytical method: the density surface of locations of economic

activity is compared against one derived from centrality indices at street junctions. For clarity, the kernel in the latter case is grid-based; that is, it represents the distribution of network centrality in 2D space (it is also possible to perform KDE with the network as the base, see Okabe *et al.*, 2009). In both cities, a positive relationship is found between economic activity and network metrics - betweenness in particular - and this suggests that network features are indeed a strong predictor of human activity. This suggests that such analysis might fruitfully be applied to the field of crime.

3.2 Empirical analysis

On the basis of the literature reviewed above, three facts are apparent: that a relationship between the street network and crime is anticipated by theory; that existing studies suffer from certain shortcomings in their measurement of the issue; and that techniques exist via which such a relationship can be meaningfully quantified. In this section, analysis is presented for the crime of residential burglary in Birmingham, UK, in which two particular features of the street network are compared against levels of offending. No previous research has used network analysis of this type in the context of crime.

Before analysing crime explicitly, analysis of Birmingham's street network will be presented. Within this, issues related to data preparation will be discussed and general trends among metrics explored. Among other things, comparison with categorical network description highlights its deficiency as a reliable proxy for network activity. After doing this, the introduction of crime data allows the relationship with burglary risk to be evaluated using a statistical model.

3.2.1 Analytical strategy

Before proceeding with the analysis, several decisions concerning the details of the analytical process are required to be made. These relate to the representation of space - essentially, the choice of unit of analysis - and the selection of appropriate

street network variables.

3.2.1.1 Units of analysis

When considering the primal representation of the street network, two fundamental units arise naturally: vertices and links (corresponding to junctions and street segments, respectively). Of these, links are an obvious choice for analysis, for several reasons. Unlike vertices, they have spatial extent and are more meaningful geographical entities, so that they can be more properly thought of as the location of crime. In a similar vein, street segments are the structures on which houses - the targets of burglary - are arranged: trivially, the location of a target is specified in terms of its street. Because of this, crimes can be meaningfully associated with segments, whereas assigning crimes to junctions on the basis of proximity has the potential to be extremely problematic. Finally, with a practical perspective in mind, street segments are a standard unit for interventions, especially those which might be deployed on the basis of network-level insight (*e.g.* directed patrolling).

Even when the street segment is taken to be the basic object, there remains the possibility of aggregating or disaggregating among these. For example, groups of segments could be considered, or individual segments could be split according to some rule. Both of these possibilities are rejected. In the first case, aggregation represents an unnecessary loss of granularity, which is at odds with one of the main themes of the research. In the second case, disaggregation is futile since the metrics in question are defined at the level of links, so that any two parts of the same segment would receive the same value. Because of its correspondence with the level of measurement, the street segment is chosen as an appropriate unit.

One alternative, and that used in the aforementioned studies of economic activity (Porta *et al.*, 2009, 2012), is to carry out analysis not on discrete network entities at all, but rather to use metrics to build a centrality ‘surface’ via standard 2-dimensional KDE. The reasons given for this are two-fold: that activities often cannot definitively be assigned to particular segments; and that certain activities

(shops, for example) can exert influence over some significant radius. In the present context, it is argued that both of these are invalid. In the first case, the quality of the data, together with the use of a text-matching algorithm, means that incidents can be determined with very high levels of confidence to have occurred on a particular street segment. In the second case, it is necessary to appeal to the difference between the two activities: whereas economic activity is measured in terms of the presence of permanent features (*e.g.* shops), crimes are point events which are highly contingent on their particular spatial circumstance. For these reasons, the analysis is based on the street segment (link) as a discrete unit of analysis.

3.2.1.2 Network metrics

The remaining choice to be made before beginning analysis concerns the selection of appropriate network metrics for which the relationship with crime can be investigated. This choice is crucial since the various metrics available have differing interpretation in terms of the aspects of network use and navigation to which they relate. From the criminological perspective, the principles in which interest ultimately lies do specify, to a certain extent, the mechanism by which urban form shapes offending. In order for inferences to be meaningful and well-aligned with these theories, then, the relationship between the two must be clearly delineated. More concisely, care must be taken to ensure that any metric chosen is a proper measure of the factors which are hypothesised to drive crime.

The criminological theories of most immediate relevance to the present study are routine activity theory and pattern theory, both of which provide the context for the majority of previous work on street network effects. In simplified terms, the primary assertion of both of these is that areas of greater use (in terms of routine activity traffic) are likely to be associated with higher crime, by virtue of the greater frequency with which they are considered as criminal opportunities. By using a measure which reflects activity in this sense, this aspect of theory can be directly tested.

With this in mind, betweenness is a natural choice in several respects. The journeys

which are enumerated during its calculation can be regarded as precisely those which constitute routine trips: those taking place between pairs of end-points, via shortest paths. Real-world awareness spaces and opportunity surfaces, built up through the accumulation of such trips, should therefore approximately be reflected by the distribution of betweenness. In addition, radially-limited betweenness represents a well-motivated refinement, since the formation of awareness spaces is likely to be dominated by local journeys (considering, for example, the dominance of short distances in journey-to-crime data; see, for example Townsley & Sidebottom (2010) and Snook (2004)). By applying such a limit, areas identified as high-use will be those which feature in many short journeys, which are more likely to represent concentrations of pedestrian travel. In addition, one final point of note is that several of the special cases explored in previous studies are implicit when betweenness is used. *Cul-de-sac* segments, for example, will be characterised by low betweenness values, since the only trips in which they feature will be those to the end of the *cul-de-sacs* (of which there are relatively few); through roads, on the other hand, will feature in many trips. With these reasons in mind, betweenness is therefore the principal quantity by which the network will be measured.

It is also worthwhile at this stage to comment upon the sense in which betweenness is related to notions of centrality used in previous work. Many previous studies have invoked some notion of ‘permeability’ as the network property under investigation, defined to reflect the ease with which places can be accessed. The accompanying reasoning suggests that areas which are more difficult to reach, in some sense, will experience less activity and, consequently, less criminality. While the effect may ultimately be the same, this is a somewhat different principle from that being measured by betweenness. In short, the discrepancy can be reduced to the distinction between the ease with which a certain section can be traversed and the likelihood of doing so, which are not equivalent. For example, a part of the network may be well-connected by streets (close to a main road, for example), but if there is no reason to travel along them, it is unlikely to see substantial activity. Conversely, if

there is some imperative to travel along a convoluted path (*e.g.* because it connects prominent end-points) then it can experience high usage in spite of its relative lack of permeability.

Consideration of permeability is certainly well-motivated, and has particular relevance to the notion of exploratory behaviour on the part of offenders. Permeable streets are more likely to be found during such exploration (or, indeed, invite it) and may also be attractive for reasons of escape. Nevertheless, such exploratory behaviour appears to be rare (Cromwell *et al.*, 1991; Ashton *et al.*, 1998), and the focus here is on a routine activity approach in particular. In that context, the interpretation of betweenness as a probability of usage is a more direct measure of routine activity density. It should be noted, however, that notions of permeability are not entirely absent from this: since betweenness is based on shortest paths, these will tend naturally to avoid less-accessible places.

If permeability were to be measured directly, metrics such as closeness and straightness would represent appropriate means of doing so. They are not included here, however, for two reasons. The first of these is technical: both metrics are defined for vertices, rather than links, and so cannot be applied straightforwardly to street segments (the central unit of analysis here). This could be overcome by, for example, taking the values for a segment to be the average of its end-points, but this introduces uncertainty and can be particularly problematic for long segments. The second reason is more theoretical: while the link between usage and betweenness is very clear, permeability is a more nebulous concept, not entirely captured by either of the above metrics. For these reasons, together with the desire for parsimony and to focus on routine activities alone, investigation of these metrics is not included here.

One additional metric is, however, incorporated. The previous study by Johnson & Bowers (2010) included a categorisation of roads as geometrically ‘sinuous’, which

has relevance to the concept of line-of-sight as invoked by space syntax. This is a clearly-defined property, and one which is self-contained for individual segments, and so is also included here. Instead of using a categorical classification, though, a new metric is introduced in order to provide a more granular and objective measurement. This is defined as the *linearity* of the segment, denoted L , and measures the extent to which the geometry of the segment deviates from a straight line. For a given segment, e_i , it is defined as the ratio of the Euclidean (as-the-crow-flies) distance between its two end-points, s_i^{Euc} , to the true arc length of the segment curve, s_i^{arc} , so that its value is

$$L_i = \frac{s_i^{Euc}}{s_i^{arc}}. \quad (3.6)$$

A perfectly linear segment, for which the two lengths will be equal, will therefore take the value 1, with the value decreasing as the deviation from a straight line increases. The categorical system employed by Johnson & Bowers (2010) is, again, implicit within this, since segments could be classified as ‘sinuous’ if their linearity falls below some threshold value.

3.2.2 Street network

In this section, analysis of the street network of the city of Birmingham, UK, is presented. After discussing the nature of the data used and a significant volume of necessary pre-processing work, a number of basic measurements for the network are shown. While the primary motivation for this is to establish a foundation for the analysis of crime which follows, these results are of consequence in their own right both as measurements for a city which has not previously been analysed in this way, and as a demonstration of the inadequacy of administrative categories for the purpose of research of this type.

3.2.2.1 Data

The street network data used in this work was produced by the Ordnance Survey (OS). In this dataset, the network is specified on the basis of junctions, where a junction is defined as any point at which two roads intersect (in a physical sense;

that is, regardless of street name or continuity) and a street segment is any portion of road which connects two junctions. This corresponds exactly to the primal representation. The dataset is structured by street segment, so that each segment is represented by a record, which includes both its geometry and several items of contextual information. These include a hierarchical classification ('Motorway', 'A Road', 'B Road', 'Minor Road', 'Local Street' or 'Private Road'), the nature of each street segment (*e.g.* 'Single Carriageway', 'Traffic Island Link', 'Roundabout') and the name of the street. The segments used here are all those which lie either within the city of Birmingham or within a 2 kilometre buffer around its perimeter. The reason for the inclusion of this buffer area is to mitigate the edge effects to which network metrics are susceptible: the entire dataset is used when metrics are calculated, but segments lying in the buffer area are discounted in all further analysis.

All tasks related to network data were performed in Python, with the additional modules PySAL and NetworkX used for data reading and network representation respectively. All other tasks, such as geospatial processing and metric calculation, were performed using bespoke Python scripts, since the intricacy of many of these was beyond the capability of existing modules. This was also the reason for scripting manually rather than using Geographical Information Systems, such as ArcGIS.

3.2.2.2 Network pre-processing

Because of the way it is constructed, the OS dataset contains a number of undesirable features from the point of view of metric calculation. Though there are many particular cases, all are manifestations of the same general issue of superfluous features: the presence of junctions in the data which would not be considered true junctions in the real world. The clearest example of this is the issue of roundabouts, as seen in Figure 3.4a. At roundabouts, each exit is treated as an individual junction, so that the structure is represented as a series of very short segments. Although this is technically correct, it does not reflect well the function of roundabouts: from the point of view of wayfinding, for example, there is little conceptual difference between these and 'simple' junctions. The presence of the additional segments, which

is simply an artefact of the physical arrangement, does, however, significantly distort network measurements.

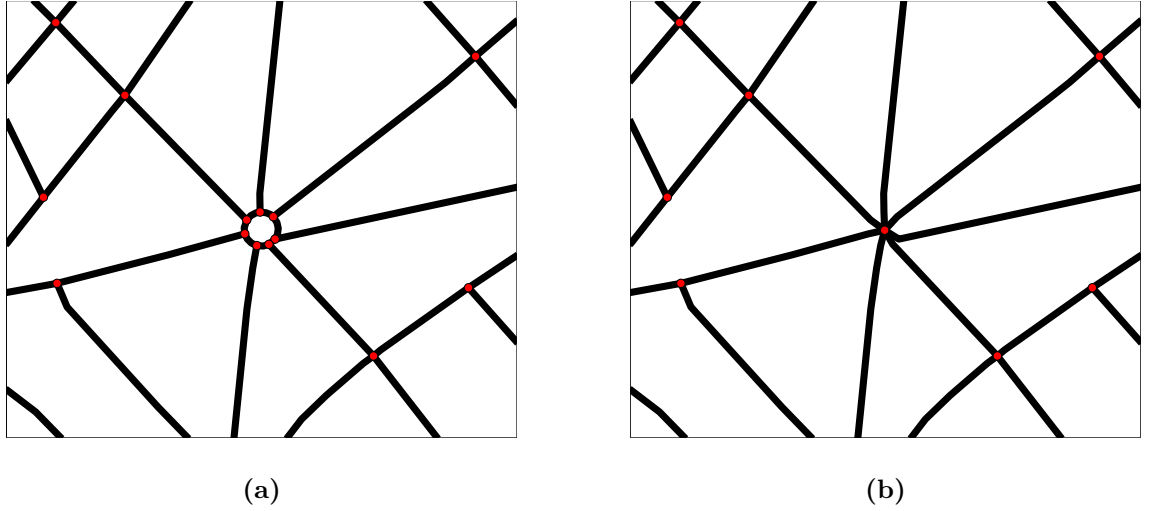


Figure 3.4: The pre-processing of roundabouts: a) the raw form of the data, in which the roundabout is represented as a series of short segments, and b) the state of the network after processing, where the roundabout has been ‘collapsed’ to a single vertex.

This distortion can be seen by considering the example of betweenness (though it also applies to other metrics). In the calculation of betweenness, all junctions are treated equally and each contributes to the calculation via the journeys which terminate at it. When one junction is represented by several vertices, though - as is the case here - that junction’s trips are effectively counted multiple times (once for each vertex) and its influence is exaggerated. The result is that the presence of a roundabout implies a false density of junctions, thereby distorting calculations. A secondary problem is also present when network distance is interpreted topologically: the number of ‘hops’ required to cross a junction, which ought properly to be only 1, is much higher, thus artificially inflating the length of any path which features a roundabout. This is clearly problematic for metric calculation.

The solution to this problem used here is to ‘collapse’ the roundabout to a single vertex, as would be the case if it were a simple junction. The result of this can be seen in Figure 3.4b: all individual exits have been replaced by a single new vertex

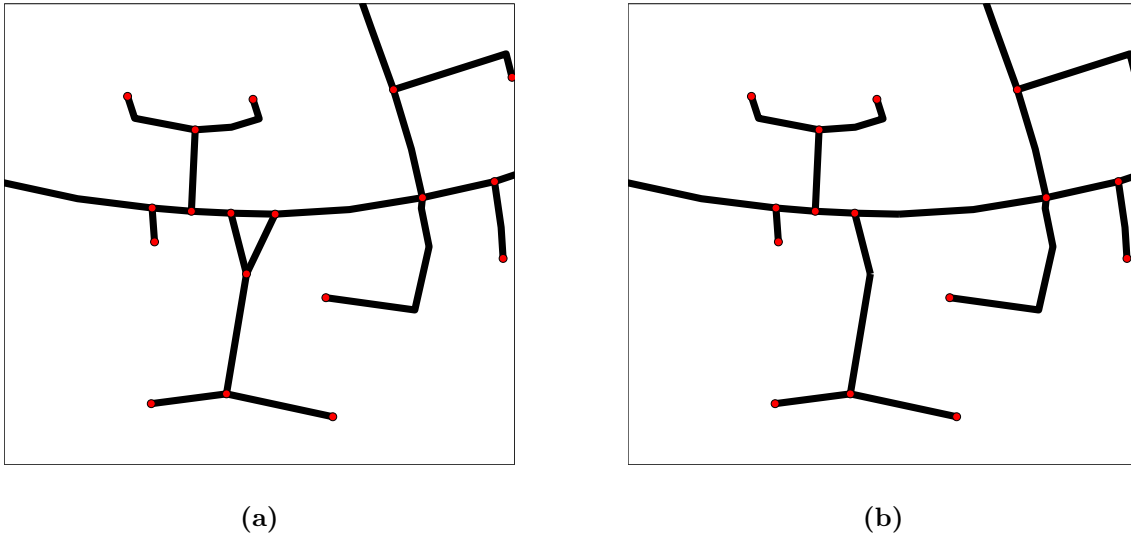


Figure 3.5: The pre-processing of Traffic Island Links At Junctions: a) the raw form of the data, in which the junction in the centre of the image is represented by 3 vertices, and b) the state of the network after processing, where one of the superfluous segments has been ‘absorbed’ into the other and 2 unnecessary vertices have been removed.

at the centre of mass of the roundabout. Configured in this way, the roundabout’s representation corresponds more closely to its role in the network.

A similar problem, though one for which the solution is different, is the presence of additional links at junctions. An example of this is the feature labelled by the OS as ‘Traffic Island Link At Junction’, which arises when a road briefly divides immediately prior to a junction. The structure, shown in Figure 3.5a, is problematic for similar reasons as roundabouts: despite representing a single junction, two additional vertices are present (one where the road divides, and one of the two where the divided road meets the other).

In this case, the solution is simply to remove one of the ‘halves’ of the divided road, along with its associated vertices; the result of this is shown in Figure 3.5b. In effect, the segment is ‘absorbed’ into the complementary half. It should be noted that, although this leads to an unrealistic physical shape, this is not a problem: the motivation for this processing is the calculation of metrics, for which the shape of streets is immaterial. Indeed, for this reason, either half of the divided road can be chosen for removal without loss of generality (as long as the topology is preserved)

and this selection is made randomly in practice.

The above point is worthy of further elaboration, in anticipation of later analysis which will require crimes to be associated with particular segments. Although the processing described here involves modification of the network, including the deletion of features, the processed form is used only for the purpose of metric calculation. In all further analysis, the raw, unprocessed, form is used. One issue to be addressed, therefore, is how metric values can be assigned to links which were absent from the network used in calculation, but are still present in the raw form. Two approaches for this are taken. In the case where one segment has been ‘absorbed’ into another (as with Traffic Island Links), the deleted segment is assigned the same value as that into which it was absorbed, since the underlying reasoning is that they are, in effect, the same segment. In the case of roundabouts, which are collapsed rather than absorbed, the segments are simply omitted from the analysis, since the occurrence of burglary on roundabouts can be assumed to be negligible.

3.2.2.3 Network properties

After processing, the network is in suitable form for the calculation of metrics, and the results of this are shown in this section. This serves the dual purpose of providing preliminary insight into the nature of the data and allowing the new metrics used here to be compared with traditional measurements.

The most basic analysis of betweenness is to consider its distribution throughout the dataset *i.e.* across the street segments of Birmingham. This distribution is shown, for various definitions of the radial limit, in Figure 3.6, and several trends are apparent upon inspection. In both metric cases, the complementary cumulative distribution function (CDF) is approximately straight-lined on semi-logarithmic axes, indicating an exponential form (this also persists for various other radii not shown). This is in agreement with published work (Crucitti *et al.*, 2006b), although it should be noted that the results here are for link betweenness, rather than vertex betweenness.

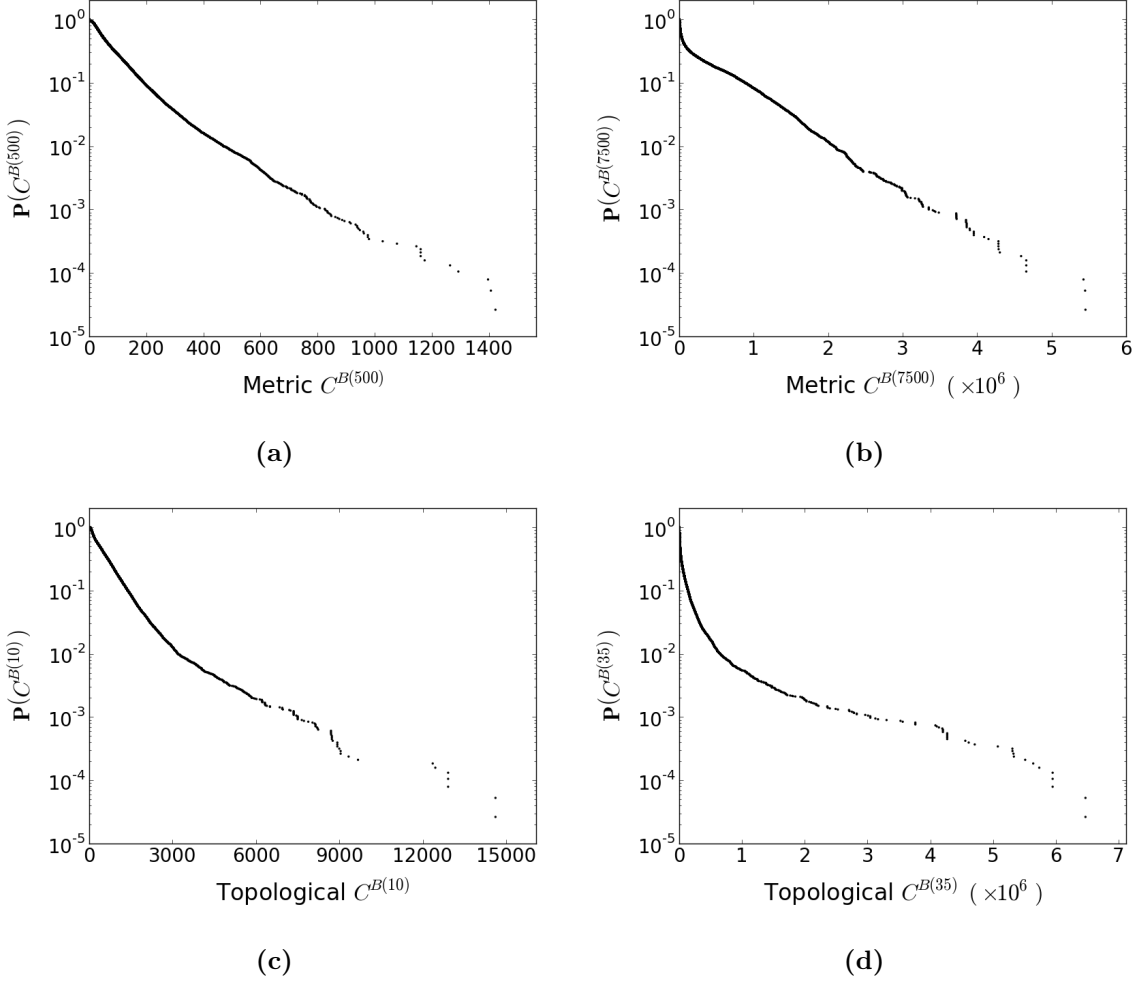


Figure 3.6: Distribution of radially-limited betweenness, $C^{B(r)}$, for the street network of Birmingham. The panels correspond to different values for the radii used: metric values of a) 500m and b) 7,500m, and topological distances of c) 10 and d) 35. Each plot shows the complementary cumulative distribution of $C^{B(r)}$; formally, $\mathbf{P}(C^{B(r)})$ is defined by $\mathbf{P}(C^{B(r)}) = \int_{C^{B(r)}}^{\infty} \frac{M(C')}{M} dC'$, where M is the total number of links and $M(C)$ the number of links with betweenness equal to C .

One other notable feature is the relatively high density of small C^B values as the radius is increased, as seen in Figure 3.6b. The reason for this is likely to relate to the manner in which a link's $C^{B(r)}$ scales with increasing r . For relatively secluded links, which do not act as ‘through roads’, $C^{B(r)}$ scales approximately with the number of vertices which lie within r , since any additional centrality as r increases is likely to be due to single trips to ‘newly-discovered’ vertices. For central links, however, along which journeys tend to be ‘funnelled’, the scaling will be at a much faster rate. This can be understood heuristically: for any new vertex included as r increases, cen-

tral links will receive a disproportionate number of the trips which terminate at it. This effect - a variant of the rich-get-richer phenomenon - explains why the growth of $C^{B(r)}$ on central streets significantly outstrips that of those with low betweenness.

The distribution is less clearly-shaped in the case of topological distance. Although the departure from straight-line behaviour is not dramatic for low radius (see Figure 3.6c), it becomes pronounced as r increases (Figure 3.6d). In addition, no clear relationship can be seen if doubly logarithmic axes are used (not shown). The reason for the very large values seen in the tail is similar to the above, but is accentuated by an additional factor. Calculating shortest paths on the basis of topological distance tends to reward links of high physical length, which allow large distances to be covered ‘cheaply’ (in terms of path length) and are therefore favoured. Links in the tail tend to be examples of this, as some become the ‘key’ links for trips between different regions of the network.

In order to show how betweenness is distributed spatially, Figure 3.7 shows maps of the street network of Birmingham, coloured according to betweenness. In these maps, the natural logarithm of betweenness is used in order to accentuate the variation at lower values and reduce the influence of the tail. Panels (a) and (b) show values for a metric radius of 3,000m, and a topological radius of 15 steps is used in (c) and (d); the higher values in the tail are clearly apparent in the latter case. In both cases, though, it can be seen how betweenness identifies a hierarchical, skeletal structure for the network as a whole. Looking at the zoomed images in particular, it appears to be successful in locating those links which would be intuitively expected to be of greatest use.

In motivating the use of betweenness, it was claimed that it is likely to be a more granular and objective measure of the use of street segments than the categorical classification provided by the OS. This can be tested by considering the relationship between classifications and C^B . Figure 3.8 shows the distribution of $C^{B(3000)}$,

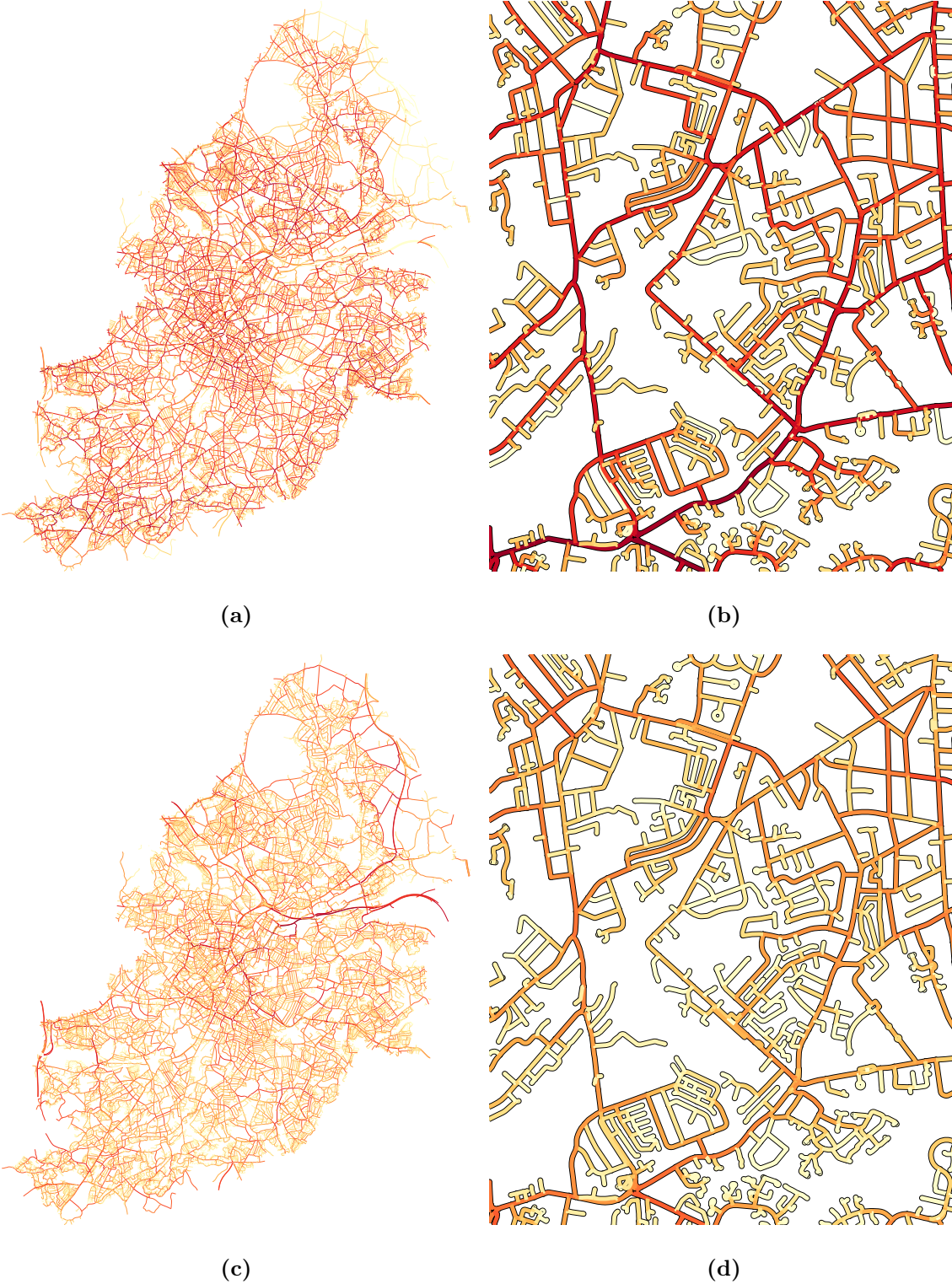


Figure 3.7: Street segments coloured according to betweenness, showing the entire street network of Birmingham, and one zoomed section. For panels (a) and (b), betweenness values are calculated on the basis of a metric radius of 3,000 metres, and for panels (c) and (d) a topological radius of 15 steps is used.

represented as a boxplot, for the various main OS classifications. Although there is indeed dominance by the more major categories of road, it is clear that the hierarchy is far from well-defined with respect to betweenness. In particular, ‘A Roads’, which would traditionally be expected to experience the highest use, do not show the highest values on average. Over and above this, however, the high variation within each of the major categories, and the considerable overlap between their boxplots, is a strong indication that these categories do not correspond reliably to centrality (in the sense in which it is measured by betweenness).

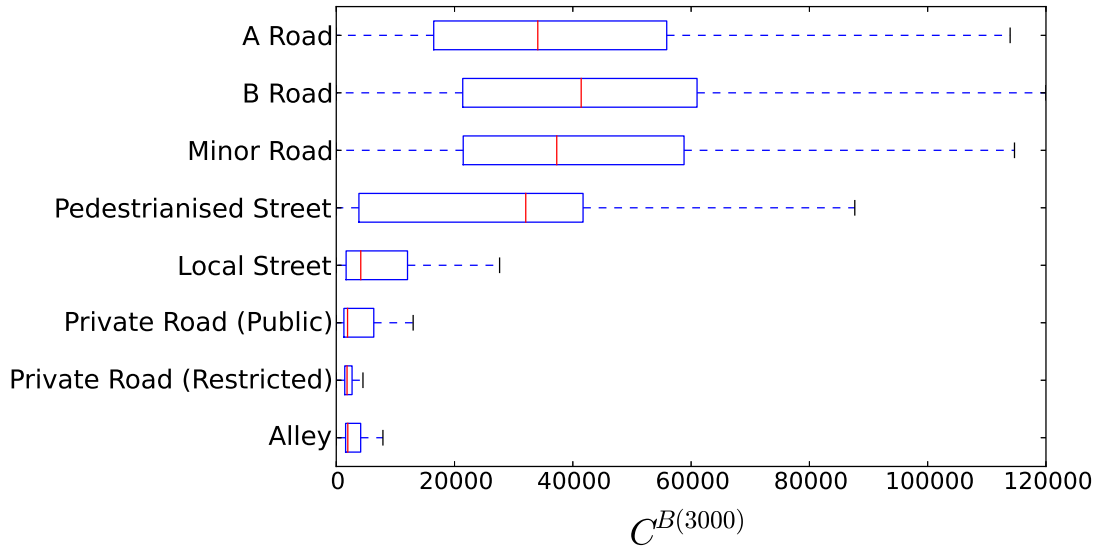


Figure 3.8: Boxplots of betweenness values, $C^B(3000)$, for various OS road classifications, using a metric radius of 3,000 metres. Categories are ordered on the basis of approximate hierarchical position. In b), the empirical cumulative distribution function for linearity, L , is shown.

The other metric which is considered in this work is *linearity*, L , as defined in equation (3.6). The distribution of L across the segments of Birmingham is shown in Figure 3.9, in which it can be seen that the distribution is dominated by large values, close to the maximum of 1. This feature, which is to be expected, is likely to limit the explanatory value of L in later analysis, but it is included nonetheless.

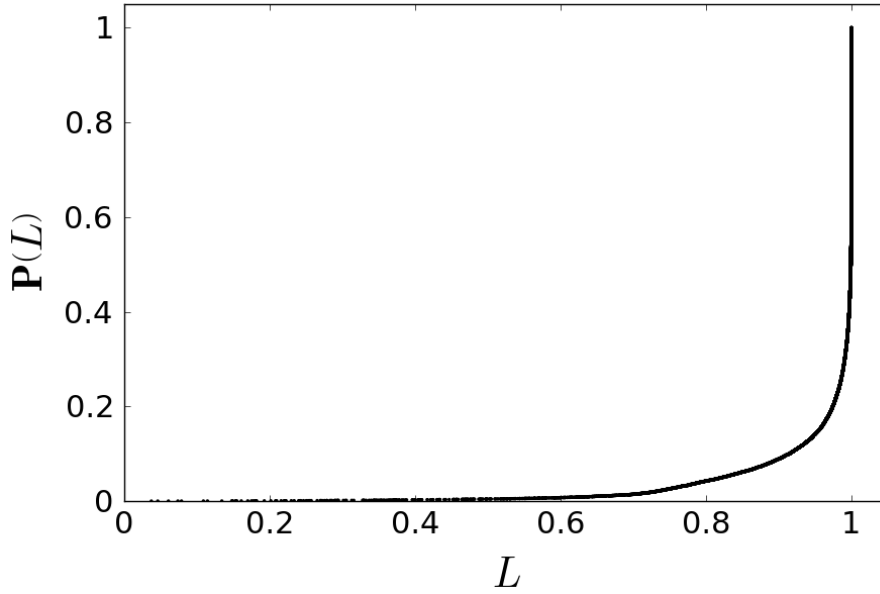


Figure 3.9: Empirical cumulative distribution function for the linearity, L , of street segments in Birmingham.

3.2.3 Crime

Having calculated various metrics for the street network, and carried out preliminary analysis on these, it remains to address the main question of this chapter: the relationship between the network and crime. This is done by considering a particular dataset, relating to residential burglary in the city of Birmingham. The first step of this analysis is to address the question of whether events are indeed distributed heterogeneously on the street network; a pre-requisite for the existence of any relationship with centrality. After this is established, the relationship with the metrics introduced in the previous section is examined via a statistical model.

3.2.3.1 Burglary data

A set of crime data was provided by West Midlands Police (WMP), comprising all recorded incidents of residential burglary for the four-year period between April 2009 and March 2013. In that period, 27,383 offences were recorded. For each record, the location of the incident is given in two forms: in British National Grid (BNG) coordinates, and as a full-text address.

In the analysis of crime levels, it is necessary to control for opportunity, in the

sense of the number of potential targets at any location; in the context of burglary, this can be understood as the number of houses present. In order to do this, the OS Address Point dataset for Birmingham, which contains the full address and BNG coordinates of every dwelling, is used.

An issue arises at this stage concerning the relationship between the crime and address datasets: it is acknowledged by WMP that the grid references associated with crimes are not accurate to the level of OS data for the corresponding address, and may not necessarily match. In addition, the full-text addresses of crimes are input by officers, and suffer from inconsistency and typographical errors. In order to ensure consistency between crime and opportunity data, therefore, an initial processing step was included for the purpose of linking each crime record to a particular OS address point (at which stage its grid reference could be updated to that given by the OS). This was done using a bespoke text-matching algorithm, whereby a search of the address file is iteratively refined on the basis of postcode, street name and flat/house number until a match is found. By this process, 26,614 incidents (97.2% of the total, with no systematic aspect to the attrition) were successfully geo-coded, so that each could be associated uniquely with an OS address point. It is these events which were subsequently analysed.

One final stage of pre-processing was required in order to associate each address point (and, therefore, crime) with the street segment on which it lies. This is, again, a non-trivial task, since a situation frequently arises whereby the segment on which a given property lies is not that to which its grid reference is closest in Euclidean space. This is caused by the grid reference being defined as the central point of the property, rather than where its entrance is situated; because of this, a ‘closest line’ match leads to errors. Although this could be tolerated in most analyses, the explicit emphasis on street metrics here, and the considerable variation possible between nearby streets, suggests that it should be addressed in order for the analysis to be robust. Again, therefore, a bespoke algorithm was used, in which address

points were preferentially associated with segments which share the correct street name. To be precise, for each address point, a 40 metre radius was searched for all segments with a matching street name: if any existed, the closest of these was taken to be the match; otherwise, the closest segment was selected, regardless of street name.

3.2.3.2 Clustering of crime on networks

As the initial step in the analysis of crime, and in order to establish the premise for the remainder of the analysis, a basic test of the hypothesis that crime is distributed heterogeneously across the network was performed. The method used was essentially a replication of that described by Johnson & Bowers (2010), in which a Monte Carlo approach was used. Analysing crime in this way, it is necessary to control for opportunity: even if crime were entirely random, some segments would experience higher levels simply because more houses are situated on them. Rather than testing whether crime is unevenly distributed over segments, therefore, the question concerns whether the distribution of crime over all opportunities is random with respect to the network. Formally, then, the null hypothesis is that this distribution is indeed random.

The disparity between the observed data and that which would be expected under a uniformly random distribution of crime can be quantified by comparing the Lorenz curve of the observed data with those derived from synthetic sets of data generated under the assumption of randomness. The computation of the Lorenz curve is simple: every street segment with at least one house is ordered according to its crime rate during the study period, and the cumulative fraction of addresses is plotted against the cumulative fraction of incidents as the ordered list is iterated through. This curve is shown for the observed data in Figure 3.10a.

In order to produce the null distribution against which this curve can be compared, 99 synthetic datasets were generated. For each of these, 26,614 homes (the number of observed incidents) were selected uniformly at random from the set of

all those in Birmingham. The sampling was done with replacement (so that a given property can be selected multiple times in the same dataset) in order to allow for repeat victimisation. Each of these represents a possible set of incidents under the assumption of random victimisation, and a Lorenz curve can be calculated for each by the method described above. The mean of these 99 curves is plotted as a second line in Figure 3.10a. The heterogeneous distribution of opportunity can be seen in the fact that this curve is not a straight line: even when victimisation is random, different segments experience it disproportionately.

The difference between observed and synthetic distributions can be quantified using the Gini coefficient (Gini, 1921), which measures the extent to which some quantity is distributed unequally in a population (in this case, incidents across street segments). It is calculated as the ratio $\frac{A}{B}$, where A is the area between the two Lorenz curves and B is the area of the graph above the curve of expected equality (in this case, the ‘Simulated’ curve). Any value above 0, therefore, indicates that the curve of the observed distribution differs from that expected under the assumption of uniform distribution.

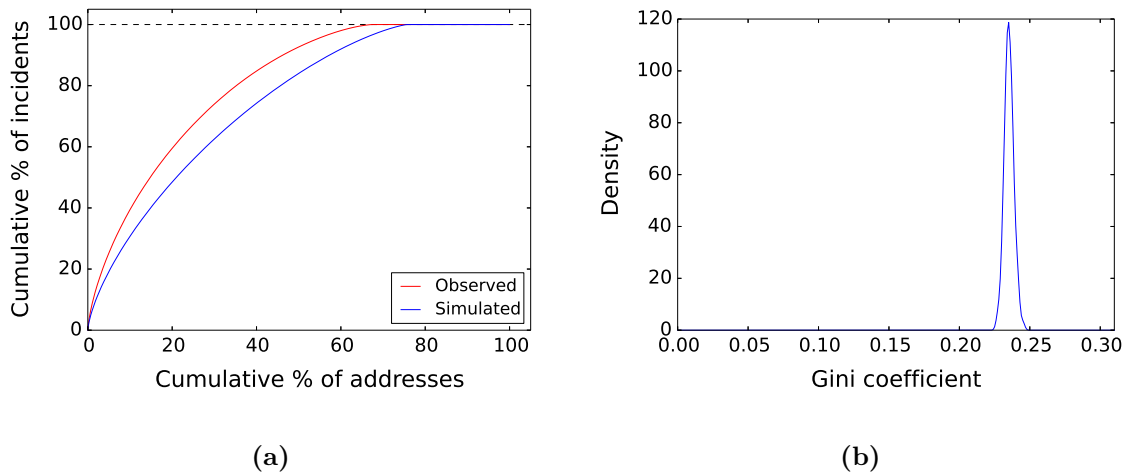


Figure 3.10: Clustering of burglary incidents at street segment level: a) Lorenz plots for both observed and simulated data, and b) the distribution of the Gini coefficient across all 99 simulated datasets.

Figure 3.10b shows the distribution of the Gini coefficient across the 99 synthetic

datasets generated. The distribution is peaked around a value of 0.24 and all density is far from the value of 0, which would be the case under the null hypothesis (if the distributions did not differ). Statistical significance can be estimated by considering the distribution of the Gini coefficient relative to this null value of 0. To do this, all 99 values of the Gini coefficient are placed in a rank-ordered list and the value R_g is taken to be the position which 0 would take in this list (North *et al.*, 2002). The statistical significance p is then defined as

$$p = \frac{n_g - R_g + 1}{n_g + 1}, \quad (3.7)$$

where n_g is the number of synthetic distributions. In this case of the test performed here, for which all Gini values are strictly positive, the hypothesis that the clustering of incidents at street segment level is due to chance can be rejected with a pseudo p -value of 0.01.

3.2.4 General relationships

Before the relationship between crime and network properties is examined statistically, an impression of the general trends present can be gained by inspecting the data graphically. Considering betweenness first, the basic hypothesis that higher C^B should be associated with higher crime suggests a simple approach whereby the average crime rate is plotted as C^B increases. To do this, all street segments with at least one dwelling are ordered on the basis of increasing C^B , and the average crime rate is plotted while moving through this ordering.

Since burglary is a rare event (and many properties are not victimised during the period of study), general trends are difficult to identify if individual segments are examined, and so the data are smoothed by considering aggregated groups of 3,000 segments. For each plot in Figure 3.11, therefore, the value plotted for a given C^B rank, R_B , is the overall burglary rate for the street segments ranked between R_B and $(R_B + 3,000)$ over the four year period.

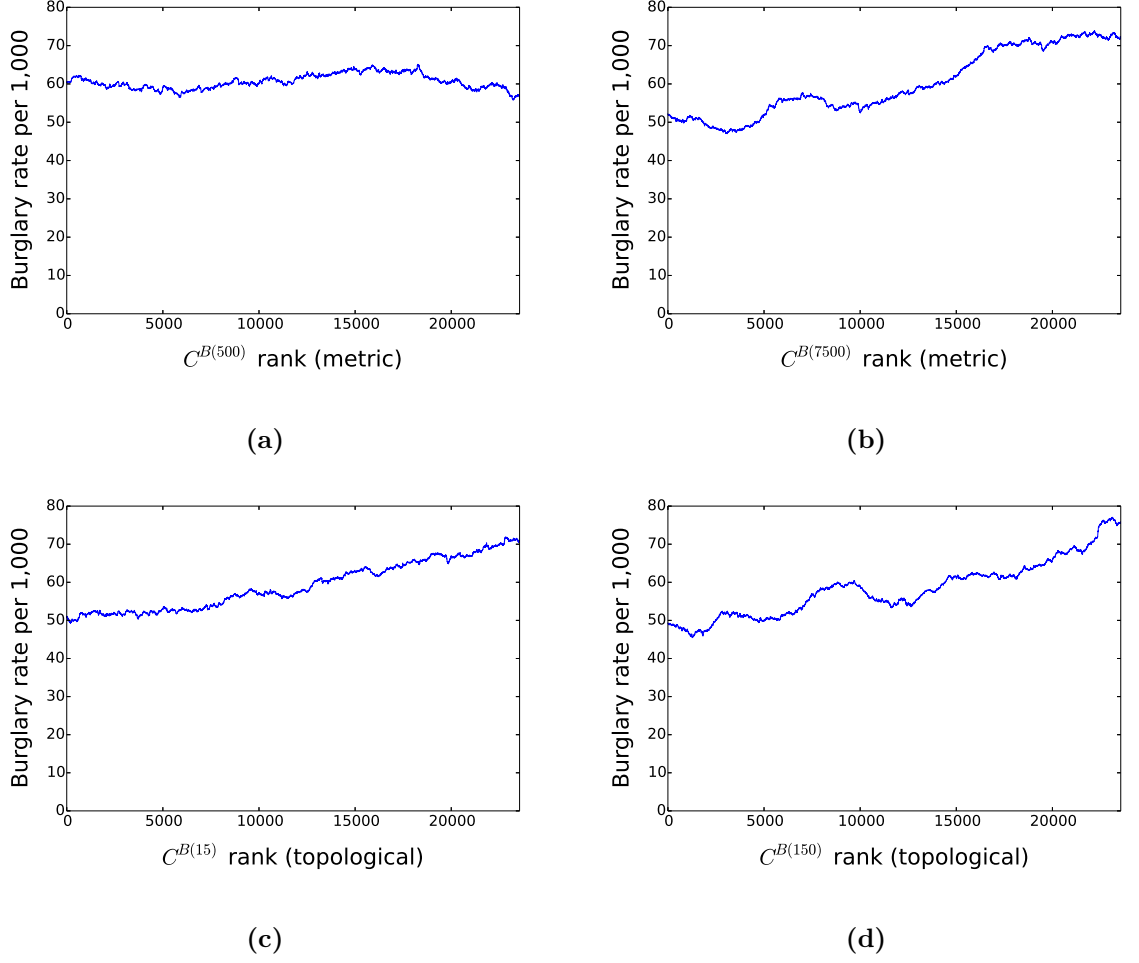


Figure 3.11: The variation of burglary rate when moving through a list of street segments ordered according to betweenness. The values plotted for any given rank represent the aggregate burglary rate for the 3,000 segments at that position in the list. The panels show the relationship for different variants of $C^{B(r)}$: for metric r of a) 500 metres, or b) 7,500 metres, and for topological radii of c) 15 or d) 150.

Overall, the figures show a positive association between burglary rate and betweenness, though it is clear that the form (and strength) of this is dependent upon the form of $C^{B(r)}$ used. When distance is measured in metric terms, the relationship only becomes clear at higher radii; in the 500m case, shown in Figure 3.11a, there is no observable trend. As the radius increases, however, a clear increasing pattern becomes apparent, and Figure 3.11b is an example of this. In the topological case, the increasing trend is evident for all radii examined. In the two cases shown, an almost-monotonic increase is seen for $r = 15$, whereas the increase, though still clear, is much less smooth for $r = 150$.

In general, the plots shown in Figure 3.11 are consistent with the hypothesis that increasing betweenness is associated with higher rates of burglary. Of course, plotting on the basis of rank in this way does not take account of the absolute value of C^B , which certainly does not increase uniformly with rank (see Figure 3.6). It does, however, correspond to the notion that betweenness defines, in some sense, a hierarchy of streets, in which a segment's role is determined by its betweenness relative to others, rather than its absolute value. Framed in this way (which might, operationally, translate to the prioritisation of streets), the relationship can be regarded as a meaningful one.

In the case of linearity, L , plotting on the basis of rank is likely to be uninformative, since the majority of segments have values very close to 1 and are therefore essentially indistinguishable. An alternative, which recalls the earlier work of Johnson & Bowers (2010), is to classify segments dichotomously on the basis of linearity and compare burglary rates within each group ('linear' *vs.* 'sinuous'). For a given threshold value, T_L , street segments are split into two groups - those for which $L_i < T_L$, and those for which $L_i \geq T_L$ - and the overall burglary rates computed for each. The values are shown, for varying T_L , in Figure 3.12.

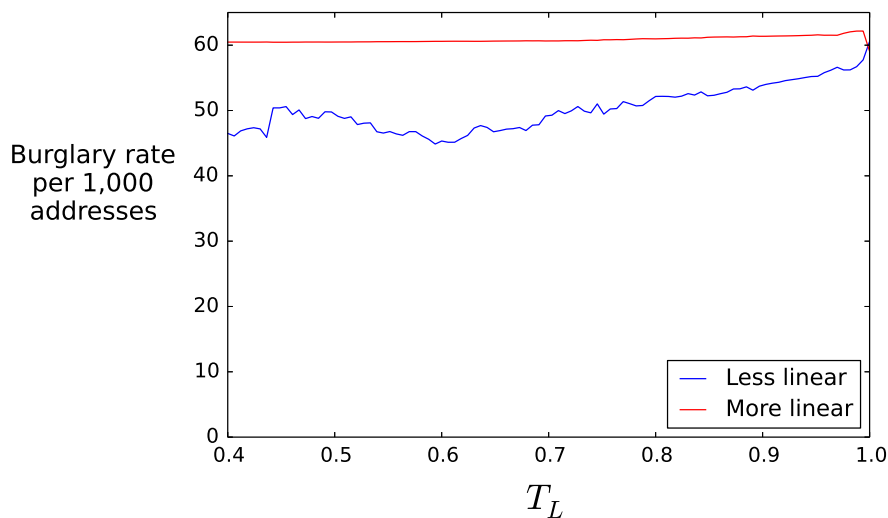


Figure 3.12: Comparison of crime rates when street segments are divided into two groups on the basis of linearity, L , falling above or below T_L .

Two general observations can be made from this: that more linear street segments experience higher rates of burglary, and that that is true across almost all definitions of what constitutes ‘more linear’. It should be borne in mind, however, that the distribution of L (see Figure 3.9) is such that the membership of these groups changes relatively little for values of T_L less than 0.9. Nevertheless, the discrepancy in rate is consistent with the conclusions of Johnson & Bowers (2010), and suggests that the linearity measure here can be used to classify streets in a similar way to the manual method used in the earlier work.

3.2.5 Regression analysis

Although the analysis presented above suggests that meaningful relationships do exist, it is, at this stage, only exploratory. Indeed, the possibility that the trends are spurious cannot be discounted: burglary patterns may be influenced by socio-demographic variables, and even in purely spatial terms it is possible that the patterns are simply artefacts of variation at some other geographical level. In order to establish the relationships rigorously, therefore, it is necessary to employ a statistical approach which controls for these factors, and this is done using a hierarchical linear model (HLM).

An HLM is a statistical object designed explicitly for application to multi-level data; that is, data which can be grouped in a meaningful and systematic way. In this case, the groupings are defined on the basis of aggregated geographical units at various spatial scales. The reason for using such a model is to account for changes in outcome caused by factors which vary at a coarser scale than the basic unit of analysis (*i.e.* the street segment). Its structure is defined formally in (3.9) below, after the various levels and variables have been introduced.

In modelling terms, area-level variation can take two forms: that which is accounted for by variables included in the model, and that which is attributable to unobserved

factors. Although the first of these can be incorporated in single-level models, they necessarily assume that the effects of the unobserved factors (modelled as random variation) do not vary systematically across units. In practice, however - and especially in a spatial setting in which similarity between nearby units is likely - this assumption is invalid. The use of an HLM addresses precisely this point.

The model considered here incorporates three levels, corresponding to three spatial units: the basic street segment and two nested census partitions. The first of these partitions is based on Output Areas (OAs), which are UK census units that contain, typically, approximately 150 households. These are taken here to correspond loosely to the notion of ‘neighbourhood’, and there are 3,127 in the city of Birmingham. These are grouped to form higher-level census geometries known as Medium Super Output Areas (MSOAs); each of these incorporates approximately 25 OAs, and there are 131 in Birmingham.

In constructing a model for burglary levels on individual street segments, it is necessary to control for opportunity *i.e.* the number of potential targets available. This can be achieved in several ways, and the choice of dependent variable for the model is contingent upon which is chosen. One option is to model the rate on each segment directly, so that opportunity is implicit by construction. The alternative is to model the raw count, and include opportunity as an explanatory variable: this can either be included in the standard way, or it can be constrained to have a certain coefficient (this is known as an ‘offset’ term).

The approach chosen here is to model counts, with an unconstrained explanatory variable for opportunity (the number of households on a segment). The reason for this choice is that potential targets are not sufficiently isolated from each other to justify their treatment as independent opportunities; there is likely to be non-linearity in the relationship between household count and opportunity. This is exemplified by the notion of ‘carrying capacity’, whereby there is some limit to the number of times

that a certain segment plausibly can be victimised, regardless of address count. The dependent variable for the model is therefore the total count of burglary across the four years of data. This quantity is modelled as being Poisson distributed, which is a standard choice when modelling count data via these methods (McCullagh & Nelder, 1993).

As well as that representing opportunity, the model comprises a number of other explanatory variables, including those related to the street network and a number of socio-demographic descriptors. At the lowest level of the model, the two network metrics introduced previously - betweenness ($C^{B(r)}$) and linearity (L) - are included, as well as the address point count and the density of properties on the segment (measured as ‘households per 100 metres’). For $C^{B(r)}$, taking the natural logarithm of the value for each segment, in order to counteract the high density of low values and large range of the distribution, was found to improve the fit of models, and it is these values which are shown. Furthermore, in order to allow meaningful comparison between variants of $C^{B(r)}$ (topological/metric, and different values of r), values in each case are normalised to the range $[0, 1]$ by dividing by the maximum value.

The variables allowed to vary at OA level are statistics from the 2001 UK census (Office for National Statistics, 2001) which correspond to factors which have been shown in previous empirical work to influence burglary occurrence. In line with research on the ages of burglary offenders (*e.g.* Farrington, 1983), the percentage of residents between the ages of 10 and 15 is included. Ethnic heterogeneity has also been shown to be associated with a significant effect (*e.g.* Hirschfield & Bowers, 1997), and is also included here. It is measured using a standard index introduced by Blau (1977), building on the work of Simpson (1949), which, for P_1, \dots, P_n representing the fractions of the population in each group, is calculated thus:

$$\text{Index} = 10 \times \left(1 - \sum_{i=1}^n P_i^2 \right) \quad (3.8)$$

In addition, the percentage of students in the area population (see Tilley *et al.*, 1999) is included: such factor is of particular relevance for Birmingham, since it is a city with several identifiable student districts. Unemployment (*e.g.* Johnson & Bowers, 2010) and the prevalence of vacant housing (*e.g.* Spelman, 1993) are further social factors thought to have an effect, and are also included. No socio-demographic variables are included at MSOA level; effects at this level are realised through the random unknown effects only.

The structure of the HLM can be written in simple mathematical terms. An indexing system is constructed for the various spatial units: i for street segments, j for OAs and k for MSOAs. The model is then fully described by

$$\begin{aligned}\pi_{ijk} &= \exp(\beta_0 + u_{jk} + v_k + \beta_1 x_{(1)ijk} + \dots + \beta_m x_{(m)ijk} + \beta_{m+1} x_{(m+1)jk} + \dots + \beta_n x_{(n)jk}) \\ u_{jk} &\sim N(0, \Omega_u) \\ v_k &\sim N(0, \Omega_v)\end{aligned}\tag{3.9}$$

where π_{ijk} is the burglary count on segment i (in OA j , in MSOA k), $x_{(1)}, \dots, x_{(m)}$ are the explanatory variables defined at segment level, and $x_{(m+1)}, \dots, x_{(n)}$ are those defined at OA level. The terms u_{jk} and v_k are the ‘random intercepts’ at OA and MSOA level respectively, both normally distributed. The results of the HLM are given in Table 3.1.

The results, as a whole, represent strong evidence that street network properties, in particular betweenness, are a highly significant predictor of burglary. In almost all cases, higher $C^{B(r)}$ is associated with higher victimisation, and the associated z-scores correspond to a very high level of significance. The one anomalous case occurs when betweenness is calculated on the basis of metric distance, with a relatively short radius of $r = 500$ metres, for which the relationship is non-significant (this is also the radius for which no discernible effect was seen in Figure 3.11). Although results are only presented for a selection of radii here, models were run for all intermediate values of r , at intervals of 500 metres, with significance shown consistently

in all but that case.

Several interpretations can be offered for the anomalous result for 500 metres. In terms of the capacity of $C^{B(r)}$ to discriminate between segments, it may be the case that there is simply not sufficient variation at this range to provide explanatory value. On the other hand, it is possible that there is a genuine absence of relationship when betweenness is calculated in that way: 500 metres is a short distance, and the meaning of betweenness may simply be too local to identify the type of throughput which might influence crime.

Although once again the result for $r = 500$ metres is anomalous, there is also a consistent significant relationship between street segment burglary rate and linearity across all other models. The direction of this relationship, however - that more linear roads experience lower victimisation - is unexpected, particularly in light of the results shown in Figure 3.12. One possible explanation for this concerns address point counts and their variation with network properties: in the data as a whole, there is a highly significant inverse relationship between linearity and address point count (so that more sinuous streets have higher address counts). The positive effect of household count on incident count predicted by the model is sub-linear at the scale of the values considered, and so the lower burglary rate on sinuous streets apparent in Figure 3.12 may simply be an artefact of their higher address count (since the calculation of rate is predicated upon an assumption of equality of opportunity). In theoretical terms, this suggests that properties are not perceived as independent opportunities, and that offenders might be attracted to sinuous roads precisely because they are known to be rich in targets.

With regard to the socio-demographic variables considered, all show significant positive association with burglary, with the exception of the percentage of vacant houses. The most striking relationships appear to be those with the size of student population and the proportion of residents aged 10-15 in an area, both of which are

supported by previous work and accord with what is known about the particular circumstance of Birmingham. Speaking of area-level effects more generally, the fact that the parameters of the random intercepts u_{jk} and v_k have been estimated to be non-zero implies that the decision to adopt the multi-level structure appears to be justified on empirical (as well as theoretical) grounds. This is, however, simultaneously a caveat: such non-zero values are indicative that a considerable amount of variation in the level of burglary cannot be accounted for by the variables considered.

	Topological			Metric		
	5	50	150	500	3000	7500
Segment level						
Address count	1.02* (146.06)	1.02* (139.88)	1.02* (141.41)	1.03* (139.83)	1.03* (158.50)	1.03* (149.12)
Addresses per 100m	1.00* (4.06)	1.00* (8.16)	1.00* (9.95)	1.00* (16.75)	1.00* (13.37)	1.00* (13.63)
Linearity	0.52* (10.65)	0.57* (9.18)	0.62* (7.76)	1.02 (0.29)	0.61* (7.95)	0.62* (7.80)
Betweenness value	4.65* (22.28)	3.86* (18.60)	3.14* (16.09)	1.00 (0.10)	2.25* (16.24)	2.27* (16.27)
OA level						
Ethnic heterogeneity	1.03* (2.85)	1.03* (2.59)	1.03* (2.83)	1.03* (2.99)	1.03* (2.84)	1.03* (2.70)
% aged 10-15	1.02* (4.04)	1.02* (4.31)	1.02* (4.38)	1.02* (4.07)	1.02* (4.42)	1.02* (4.50)
% unemployed	1.01* (2.35)	1.01* (2.57)	1.01* (2.56)	1.01* (2.34)	1.01* (2.29)	1.01* (2.38)
% students	1.01* (4.90)	1.01* (4.97)	1.01* (4.95)	1.01* (5.12)	1.01* (5.16)	1.01* (5.04)
% houses vacant	1.00 (0.25)	1.00 (0.33)	1.00 (0.33)	1.00 (0.55)	1.00 (0.28)	1.00 (0.27)
Random intercepts						
OA random variance	0.26* (24.83)	0.25* (24.76)	0.25* (24.76)	0.26* (24.92)	0.26* (24.84)	0.26* (24.85)
MSEA random variance	0.09* (6.82)	0.09* (6.80)	0.09* (6.82)	0.09* (6.81)	0.09* (6.79)	0.09* (6.78)

Table 3.1: The results of several HLMs, where the various columns correspond to different definitions of betweenness (in terms of the radius used in its calculation and whether distance is defined in topological or metric terms). The values given are exponentiated regression coefficients, and therefore represent the factor by which the mean count on a segment is estimated to change as a result of a one-unit change in the explanatory variable. *Z*-scores are given in brackets, and * denotes significance at $p=0.05$ level.

3.3 Discussion

That street networks should play a role in shaping patterns of crime is an immediate corollary of various theories of environmental criminology, such as pattern theory, though the manner in which such an effect is manifested is the subject of much conjecture. Numerous empirical studies on the topic have been carried out; however, although the conclusions of these are (largely) consistent - that more prominent or well-connected streets experience higher crime - the senses in which these terms are defined varies significantly. Moreover, the way in which these characteristics are measured is unsatisfactory in several respects: the quantities investigated suffer variously from a lack of granularity and objectivity, and their relationship to theory is often oblique. The aim of this chapter was to formalise the study of this topic by applying techniques of network science, by which network properties can be studied in a more rigorous and well-motivated way. The relationship between these and the occurrence of burglary was investigated here using the example of Birmingham.

3.3.1 Relationships established

The primary network metric investigated here was betweenness, which estimates the extent to which a given network element features in journeys through the network. When applied to street networks, it can be regarded as a proxy for the likely throughput on any given street segment when traffic is flowing on the network. As such, and in contrast to metrics used elsewhere, it is a direct and objective estimate of street network usage, which is fundamentally related to environmental theories of crime via the notion of urban awareness space. In addition, its usage is also beneficial because of the level of granularity it affords, providing a level of detail significantly beyond, for example, a categorical approach.

These issues were demonstrated in the first stage of analysis, in which the street network of Birmingham was considered in isolation. Comparison between administrative classifications and betweenness exemplifies one of the key advantages of the

approach used here, and illustrates the extent to which it differs from those used previously. The large within-class variation when betweenness is measured by street type, and the significant overlaps between classes, imply that street classifications carry relatively little information and define a far-from-consistent hierarchy in material terms. Betweenness, on the other hand, offers high resolution and appears to reflect intuitive notions of centrality when shown graphically.

The statistical model presented here indicates a highly significant positive relationship between burglary and betweenness centrality. Because of the way in which betweenness is calculated, this has an immediate real-world interpretation: that those street segments which are more frequently used during regular movement around the network will be at higher risk of burglary. Such a relationship is predicted by crime pattern theory, and these results therefore provide strong evidence that the influence of the urban backcloth on crime distribution is consistent with those principles. Insofar as betweenness can be taken to represent the broader notion of ‘centrality’, these results are also in agreement with previous work using different metrics.

The result for betweenness is subject to one caveat, related to variation in the precise way in which it is defined. Although the results were generally robust to changes in the way that shortest paths were calculated (topological/metric and for varying maximum trip length), one anomalous non-significant result was found for 500 metre radius. That particular result might be explained by the fact that the effects measured in that case are simply too local to influence burglary; however, it also highlights the fact that variations in these parameters do indeed influence results.

Concerning the case of linearity, the results consistently show that, when measured in this way, more sinuous street segments are at higher risk of burglary after controlling for other variables. Although this contradicts some previous work, the theoretical context in this regard is sufficiently diverse that the result can still be

rationalised. One possible inference may be that the ‘cover’ afforded by sinuous streets is favourable for crime and that, equivalently, the visual guardianship on more linear streets is indeed a substantial impediment.

Caution should, however, be exercised when interpreting these results and determining the suitability of linearity as a metric. As defined, it does not necessarily correspond to a street’s visual properties as, perhaps, an angular measurement would, and it is also measured separately for each segment, taking no account of visibility from adjoining segments. Coupled with the fact that linearity also appears to be associated with higher betweenness and lower address counts, the picture in this case is not necessarily clear and merits further investigation.

3.3.2 Practical implications

The relationships observed here are of relevance to several aspects of policy and practice relating to crime reduction. Most obviously, they provide cause for the addition of betweenness to the array of factors used by police in directing general policing effort and targeting specific interventions. In short, centrality represents another means by which the inherent level of risk to an area can be evaluated. In fact, network metrics may be of particular value in this sense since they are objective quantities which can be measured without consideration of any local context; they can be calculated simply on the basis of a map. As such, they are not subject to any of the inherent unreliability of other predictors, such as socio-demographic factors, and can be calculated in real-time, being updated as changes - temporary or otherwise - occur.

Implications can also be considered at a somewhat higher level of policy. Betweenness is determined by the way in which the street network is constructed, which is fundamentally an issue of design and planning. Changes to design principles can therefore cause variation in the values which are taken: the extreme examples of grid-like systems (where all streets have similar properties) and tree-like structures

(which are hierarchical) are convenient illustrations of this. Perhaps, then, it is possible to better design urban areas so as to reduce crime, with an obvious opportunity being to avoid the occurrence of high-betweenness segments.

Whether this is possible depends crucially, however, on whether the variation of crime attributable to heterogeneity of betweenness would simply be redistributed under such a change. To simplify, if it is assumed that all else is equal, it may be the case that the total volume of crime is constant and that the effect of inter-segment variation is simply to create inequality in its distribution. In this case, where the effect of a more homogeneous structure would be to distribute the same volume of crime more evenly, such a design would be of questionable value (indeed, since it removes the possibility of targeted policing, it may be disadvantageous).

On the other hand, if the additional crimes on more central streets occur, in some sense, *because* the street is more central, then avoiding such imbalances may well have the potential to reduce crime. The issue is essentially one of displacement, or rather placement (see Barr & Pease, 1990), since any notion of intervention is in the design of the city itself. Any related hypothesis would be difficult to test, since any comparison between urban areas with differing structure would be compromised, to a certain extent, by the presence of variation in many other relevant factors. An appealing study, however, might consider temporary modification of part of a street network, such as by road closure. Since even small such changes can cause large perturbations in network metrics, this would represent a direct test of the effect of centrality. The use of road closure as a general crime-prevention tactic has previously been reviewed by Clarke (2004), who refers to a number of specific deployments in several countries. Results appear to be generally positive, although there is a paucity of rigorously-evaluated interventions and few which target burglary in particular.

3.3.3 Implications for modelling

The results also have significance for the development of models of crime. The relationships identified - not only those for network metrics, but also the more basic finding that crime is clustered on networks - underline the fact that street networks are structures which play a crucial role in shaping patterns of crime. In terms of understanding why crime occurs in some places and not others, it is clear that networks represent a significant constraint and therefore ought to be taken into account; nevertheless, this is rarely the case. This point provides the primary motivation for the work in Chapter 5, which concerns the development of a crime model explicitly situated on a street network.

Network metrics, such as betweenness, can also be expected to feature in such a model. All models feature some quantity representing the underlying risk of burglary at some location and, as noted above, network metrics can be considered to influence this. In a prospective model, therefore, the effect of the network is not only manifested via the substrate on which dynamics occur, but might also be represented by including terms in the dynamical equations themselves. Again, this will be a feature of the model developed in Chapter 5.

3.3.4 Further work

The findings of this chapter are sufficient to establish the importance of concepts such as betweenness, and to provide motivation for the further analysis and modelling which follows. Remaining within the scope of this chapter, however, there are a number of possible enhancements which are worthy of further investigation in future. Within the definition of betweenness, for example, there is potential for further refinement in order to incorporate additional aspects of urban form. For example, journeys to and from known centres of activity (*e.g.* dense residential areas, entertainment districts) could be assigned higher weighting, and more sophisticated heuristics for route choice could be used. Of course, these changes would have a cost in terms of parsimony and universality.

As a final point, these results are based on data for one crime type, in one city, and have been established in general terms. Extending the analysis to other crimes and circumstances would be valuable in terms of establishing external validity. In addition, it may be of interest to examine possible interaction effects between network properties and other characteristics *e.g.* time-of-day and other features of urban form.

Chapter 4

Network effects in repeat victimisation

On the basis of the results presented in Chapter 3, it can be said with some confidence that a significant relationship exists between properties of the street network and the long-term risk of burglary. In this chapter, the analysis is extended to also consider short-term effects; in particular, the phenomenon of (near-)repeat burglary victimisation. It will be argued that, according to criminological theory, street network effects might be expected to influence the locations of follow-up incidents, relative to the ‘trigger’ incidents which preceded them. In order to measure this, a novel network metric - referred to as *commonality* - is proposed: this measures the extent to which pairs of links co-occur in paths through the network. Using a discrete choice model of target selection, this is shown to be a statistically significant predictor of the location of follow-up incidents. This implies that (near-)repeat effects are directional - they do not act uniformly in all directions - in contrast to the assumptions of many existing models.

4.1 Background

In Chapter 3, the concept of ‘risk’ was explored in a static sense, in that no account is taken of the temporal aspect of crime: the results simply concern aggregated patterns over an extended time period. While this is valuable, empirical research has demonstrated consistently that considering the spatial component of crime in isolation affords only a partial understanding of patterns of offending (Johnson & Bowers, 2004a,b). Perhaps most importantly, patterns revealed by considering the

time-course of victimisation suggest that it is possible, to an extent, to predict dynamically the locations at which crimes might occur over short time-scales and, more generally, to model the evolution of patterns of crime (Bowers *et al.*, 2004).

The patterns revealed when time is considered in conjunction with space are, in general, variations on the concept of space-time clustering, whereby incidents occurring close in space tend also to occur close in time. This is the same general phenomenon as was investigated in Chapter 2, in which ‘event networks’ were used to characterise different forms of clustering. In that chapter, clustering was considered in relatively abstract terms, in the sense that it was regarded simply as a phenomenon to be characterised and described without detailed recourse to its relationship with theory. When such patterns are considered in more specific contexts, however, their nature and the possible reasons for their formation can be described in more concrete terms.

Urban crime, and burglary in particular, is one example for which this is the case. In this setting, clustering is most obviously manifested through the phenomena of repeat and near-repeat victimisation (see Pease, 1998; Morgan, 2001). The meaning of these concepts can be defined in several ways (repeated offending against the same individual, or group of individuals, for example) but, when framed in spatio-temporal terms, they refer to cases where an offence at a particular location is followed soon after by another at the same location, or elsewhere in the near vicinity. The prevalence of this phenomena has been demonstrated for numerous crimes (Pease, 1998; Grubestic & Mack, 2008), and appears to be ubiquitous for burglary in particular (Johnson *et al.*, 2007).

A number of hypotheses have been advanced as possible explanations for why offences should occur in this way, incorporating such notions as risk heterogeneity and event dependence (described in detail in the following section). These hypotheses are, in turn, grounded in more general environmental theories of crime, such

as routine activity theory, pattern theory and the optimal forager principle. It is immediately apparent that many of these are the same as those invoked when motivating the study of the street network in the previous chapter: ‘awareness space’, of which the street network is a primary determinant, is a central concept in both cases. The existence of this common theoretical foundation suggests, therefore, that network effects might be expected to also play a role in space-time clustering.

In some respects, the manner in which the influence of the network might be exerted is merely an extension of the static case: to the extent that it is described by risk heterogeneity, space-time clustering can simply be seen as a secondary consequence of inherent potential for crime on some streets. When other aspects are taken into account, however, a change of perspective is required. Whereas static patterns can be considered globally - in the sense that only the absolute location of an incident is important - understanding repeat victimisation necessarily implies a local perspective. This is because the conceptual meaning of any (near-)repeat offence is defined by its location *relative to* the initial incident: it is precisely that link which defines it as a repeat. If the victimisation is thought of as the outcome of a choice process - either explicitly, as is the case under a same-offender hypothesis, or otherwise - then the choice is made, in some sense, from the perspective of the initial incident.

This change of perspective represents a technical challenge for the analysis and modelling of street network effects. Study of (near-)repeat victimisation concerns two primary issues: whether a given incident is followed by a (near-)repeat and, if so, the location of the follow-up incident. For the second of these, the fact that location is defined in relative terms suggests that streets should be considered in terms of their characteristics relative to the initially-targeted street. Few existing network metrics provide this form of comparison, and to do so therefore requires the development of new methods for network measurement. A possible candidate, referred to here as ‘commonality’, will be defined and used in the analysis.

The purpose of this chapter is, therefore, to explore the relationship between street networks and repeat victimisation, thereby extending the approach of Chapter 3 to apply to dynamic patterns. As with that chapter, the example of burglary will be used.

The existence of such a relationship would have significant implications both for the modelling of crime and for practical interventions. While the concept of (near-)repeat victimisation *per se* is relatively well-understood, there has been relatively little research concerning whether there is heterogeneity of risk among possible targets for follow-up incidents. Accordingly, risk elevation is typically modelled as being isotropic, simply acting uniformly in all directions around an initial incident. If the street network can be found to influence the location of secondary victimisation, this would provide a means of inferring directionality in the elevation of risk, thereby allowing preventative efforts to be directed more precisely.

The chapter will begin by presenting the theoretical background concerning space-time clustering in urban crime, particularly in the context of repeat and near-repeat victimisation. This will then be used to motivate the study of network effects, before the hypothesised effects are tested using burglary data from Birmingham, UK. In the first instance, it will be shown how an existing test of space-time clustering can be adapted to the network setting, demonstrating that clustering is indeed still present when space is considered in these terms. The focus will then move to testing for directionality in patterns of near-repeat offending. The concept of ‘commonality’ will be introduced and measured for the network of Birmingham, before incorporating it in a discrete choice model of target choice in near-repeat offending.

4.2 Theoretical background

In studies of crime which explore its concentration in space, analysis is typically performed at some level of temporal aggregation. Often, this is necessary in order for meaningful patterns to be reliably identified, but it also relates to the practical

aim of identifying trends which are, at some scale, persistent. Incidents can be aggregated at several scales (*e.g.* hours, days or weeks) and any relationships which are identified apply across the period in question as a whole. Previous research related to the identification of places of particular criminal character (*e.g.* Block & Block, 1995; Sherman *et al.*, 1989; Brantingham & Brantingham, 1995) is of this type, as was the work presented in Chapter 3.

For crime patterns identified in this way, though, significant fluctuation in crime levels may be present within the temporal window considered. This can take the form of daily cycles, perhaps corresponding to the rhythm of regular activities, or as seasonal variation, for example. In any such case, it is possible that the aggregated trend is not necessarily valid at all points within the period, since the spatial concentration can vary significantly over short time-scales: several studies have found, for example, that hot-spots of crime can be transient or unstable (Barr & Pease, 1990; Weisburd *et al.*, 2004). In work by Johnson & Bowers (2004b), hot-spots are characterised as being ‘slippery’, in the sense that their location can shift or drift over time (see also Johnson *et al.*, 2008).

That hot-spots display such properties implies that the interest arising from the temporal component of crime is not restricted to simple temporal clustering, but that an interaction exists between temporal and spatial distributions. In particular, there is a tendency for incidents which occur close in time to also occur close in space, giving rise to observable space-time clusters (Townesley *et al.*, 2003; Johnson *et al.*, 2007; Grubestic & Mack, 2008).

In the case of burglary, this is exemplified by the phenomenon of repeat victimisation (reviewed by Farrell, 2005), whereby victimised properties, for a period after the initial event, are subject to a rate of further victimisation greater than that which would be expected by chance. Indeed, the temporal component of this is particularly distinct, with risk appearing to decay exponentially with time after an

initial event (Johnson *et al.*, 1997). Going further, the concept can be extended to that of near-repeats (Morgan, 2001), whereby properties close to an initial event also experience elevated risk for some period afterwards. These phenomena, which act as significant drivers of burglary hot-spots in urban environments, appear to be ubiquitous in data from several countries (Townsend *et al.*, 2003; Johnson *et al.*, 2007). In addition, since the existence of such patterns implies that some crimes are, to a certain extent, predictable, phenomena of this type represent an attractive opportunity for crime prevention (Everson & Pease, 2001).

The empirical observation of the near-repeat phenomenon implies that burglary victimisation cannot be understood by considering properties in isolation, and that they must instead be considered in the context of their wider neighbourhood and the activity occurring within it. In attempting to account for clustering of this type, much attention has focussed on two (non-mutually-exclusive) hypotheses - *risk heterogeneity* and *event dependence*, also known as the *flag* and *boost* accounts, respectively - both of which seek to provide explanations based on factors acting at the level of individual properties (Pease, 1998).

The fundamental line of reasoning of the flag hypothesis is that repeat victimisation can be explained as a statistical by-product of regular offending. It is based on the fact that, within any area, the time-stable risk of burglary at individual properties will vary considerably, and that this variation gives rise to patterns which deviate from what would be expected on the basis of chance, when the area is taken as a whole. The term 'time-stable risk' here refers to any non-varying factor which may influence the probability of victimisation, such as the type of property, presence of security features, affluence or location.

That this can give rise to levels of repeat victimisation which are inconsistent with a random process can be understood via a simple argument. For any given property, some repeat victimisation would be expected to occur by chance, at a rate

determined by the inherent risk at that property. If all properties had equal risk, the patterns arising from this would not be statistically significant when compared against a simple Poisson process. If, however, risk varies across properties, then the observed separation between incidents is a mixture of that generated by relatively high-risk properties and that generated by low-risk ones. This biases the process - in effect, the observed separations are being compared against a false base rate - and gives the illusion that the pattern cannot be occurring due to chance. This argument can easily be extended to near-repeats by observing that nearby properties are likely to be at similar risk.

The other explanation invoked in explanations of (near-)repeat victimisation is the boost hypothesis, which states that, for some period after an initial event, the risk to nearby properties is temporarily elevated. Why this should be the case can be explained by appealing to theories concerning the behaviour of offenders; in particular, the optimal forager principle (Johnson *et al.*, 2009b). This is an ecologically-inspired analogy which asserts that offenders operate in a way which is guided by the desire to increase resources while minimising the effort expended, and risk undertaken, in doing so. In terms of target choice, it implies that they will choose properties with higher anticipated rewards while being constrained by such factors as the distance they are prepared to travel and the prospect of detection.

The relevance of this line of reasoning to (near-)repeat victimisation becomes apparent when considering the situation facing an offender after the successful commission of a burglary. In committing the first offence, the burglar acquires a certain amount of relevant knowledge: the layout and means of entry to the property, the potential rewards available, and the level of surveillance. This can inform the choice of future targets: if, and when, the offender decides to commit another offence, he or she has a fundamental initial choice between the known property and other possible targets, about which at least some of these facts are unknown (see Farrell *et al.*, 1995; Ashton *et al.*, 1998; Pease, 1998). In line with the risk-minimisation principle, therefore, the

previously-victimised represents a ‘safe option’, in some sense, and therefore is more likely to be targeted.

This reasoning also extends to nearby properties and to the near-repeat phenomenon. Nearby houses are likely to be similar to the initially-victimised house (in their affluence, for example), so that the knowledge gained during the offence applies, to some extent, to them also. In terms of foraging, the fertility of one location provides an indication of the fertility of its surroundings and therefore contributes to any cost/benefit calculation. At an even more basic level, though, one aspect of fertility is simply awareness: the first stage in the appraisal of any potential target is knowing of its existence. The fact that an offender has targeted one property implies that he or she has some awareness of the local area: it may be part of the offender’s routine activity space, or it may have been encountered in the journey to or from the initial crime. Quite apart from their intrinsic appeal, this alone might be sufficient to bias a target choice process towards other properties in that nearby area.

The reference to routine activity space in this argument is an indication that various wider theories of environmental criminology, such as pattern theory (Brantingham & Brantingham, 1993a) and routine activity theory (Cohen & Felson, 1979), are again applicable here. Though the mechanism is somewhat different, the underlying reasoning, that the risk to properties can be partly explained by the extent to which they are encountered by potential offenders, is fundamental to the issue of (near-)repeat victimisation as well as to victimisation *per se*. As argued in Chapter 3, this implies that the street network can be expected to play a role in shaping the patterns of such offending. Though it is unnecessary to rehearse the argument presented there, the street network is a primary determinant of urban activity patterns and can therefore be used to estimate the factors which are predicted by theory to play a role in victimisation. These ideas will be explored further in Section 4.3.3.1.

One point which must be borne in mind when evaluating the boost account is that it

relies on an assumption that the (near-)repeat pairs of offences which it seeks to explain are committed by the same offender. Although this cannot definitively be said to be the case, it is an assumption that is lent credence by interviews with offenders (Cromwell *et al.*, 1991; Ashton *et al.*, 1998; Summers *et al.*, 2010). In addition, examination of police detection data has also suggested that (near-)repeat incidents can, in the majority of cases, be identified with the same offender. Bernasco (2008), for example, used data from The Hague, Netherlands, to show that 77% of burglaries occurring within 200 metres and 15 days of each other involved the same offender. Similarly, Johnson *et al.* (2009b) found same-offender involvement in 76% of near-repeat burglaries occurring within 100 metres and 14 days in Bournemouth, UK. These studies provide strong support for the association of near-repeat offences with a common offender.

It is also worth noting, though, that much of the reasoning concerning the boost account is still valid even if the same-offender assumption is partially relaxed. Even allowing for a different second offender, it remains the case that the second offender must have gained knowledge of the initial crime (through social contact, for example): from this point, the boost account still relies on the second offender having knowledge of the second target (either his/herself, or having been passed on by the initial offender), and so activity and awareness spaces remain crucial.

It is also important to recognise that neither the flag account nor the boost account seeks to explain (near-)repeat victimisation in its entirety. In either case, the fact that some offences can be ascribed to the effect justifies its analysis. Several attempts have been made to measure their relative contributions, including simulation-based (Johnson, 2008) and statistical approaches (Tseloni & Pease, 2003; Short *et al.*, 2009). In all cases, evidence was found in support of both hypotheses, and it is likely that both are required to generate the patterns observed empirically.

4.3 Empirical analysis

On the basis of the principles and reasoning outlined in the previous section, there is cause to expect that patterns of (near-)repeat victimisation are influenced, to some extent, by the configuration of the street network. One of the ways in which the effect may be exerted - by shaping the heterogeneity of time-stable risk - has already been established in Chapter 3; however, the issue of space-time clustering was not considered at that stage. In this section, the same data for burglary in Birmingham, UK, will be used to test for clustering and for the anticipated influence of the network.

The first stage of analysis will be focussed on establishing the existence of space-time clustering within the data; a pre-requisite of all subsequent analysis. Although there exist several means of measuring this (one of which will be used), the analysis has not previously been carried out in a network context; *i.e.* where the spatial component is expressed in terms of the network. This requires a straightforward adaptation of existing methods, and this will be shown.

Once clustering has been established, the focus moves to the primary topic of interest: directionality in the spread of burglary risk. Ideas from routine activity theory and the optimal forager principle will be used to motivate ‘commonality’, a novel network measurement which quantifies the relationships between network features in terms of travel through the network. After demonstrating its real-world meaning using examples from the Birmingham street network, its role in near-repeat victimisation will be examined using a discrete choice model.

4.3.1 Data and network representation

This section builds on the work of Chapter 3, and is conducted within the same analytical framework. The dataset examined is therefore exactly as described in Section 3.2.2.1; that is, the 26,614 incidents of residential burglary which occurred

in Birmingham between April 2009 and March 2013 and which could be successfully geo-coded to an Ordnance Survey address point.

Since the present analysis also examines the temporal aspect of victimisation, the time of each incident is also required to be known. Burglary is an ‘aoristic’ crime (*i.e.* there is uncertainty regarding the exact time at which an incident took place) and, accordingly, three pieces of temporal information are provided in the data: the date on which the incident was reported, and the earliest and last times at which the incident could have taken place. Several suggestions have been made in the literature for how best to deal with aoristic data (Ratcliffe, 2002; Ashby & Bowers, 2013), but in this analysis the report date is simply taken as a point estimate of the time of offence, for two reasons.

Firstly, some alternative approaches may cause attrition in the data (an incident may be discarded, for example, because the range of its temporal window is too large), whereas the accuracy of the recording date is consistent across all cases. From a more practical perspective, the recording date represents the time at which the police become aware of an incident, and is therefore the first point at which they may conceivably intervene to prevent a (near-)repeat. Pairs of incidents defined in this way are therefore those which might plausibly be predicted (or, at least, those which occur within the scope of a potential intervention). Aside from these concerns, sophisticated treatment of the aoristic issue is not the focus of the present work, and use of such a simple option is beneficial in terms of clarity. The majority of previous studies of space-time clustering (Johnson *et al.*, 2007; Grubestic & Mack, 2008) involve the use of point estimates, and the time of report in particular is used by Townsley *et al.* (2000), for example.

The street network used in this study is based on the same data as that used in 3, and processed as described in Section 3.2.2.2. Again, it is represented in ‘primal’ form, with the reasons for doing so exactly the same as outlined previously.

4.3.2 Clustering

The existence of space-time clustering is the fact upon which all further research is predicated; indeed, the achievability of short-term crime modelling in general depends on it. In the case of burglary, it has been demonstrated across a wide range of settings (Johnson *et al.*, 2007) using a simple statistical method. In order to establish this basic property, such analysis is repeated here.

4.3.2.1 Euclidean distance

The Knox (1964) test for space-time clustering has been applied frequently in the context of crime, and recent results for the case of burglary have been found using a version employing Monte-Carlo simulation (see, for example, Johnson *et al.*, 2007). This test was introduced, and defined formally, in Section 2.3.2; however, a more sophisticated form is used here and so the general form will be presented again.

The essence of the Knox test lies in the pairwise comparison of incidents, and the classification of each pair according to the proximity of the incidents in space and time. Prior to the analysis, an ordered system of bands is defined for both space (*e.g.* 0-100m, 101-200m...) and time (*e.g.* 0-6 days, 7-13 days...), which define various degrees of the concept of ‘closeness’ in each dimension. The incidents are then examined by comparing all possible pairs and calculating the spatial and temporal separation for each. These comparisons are used to populate a contingency table of the spatial and temporal bands (of the same shape as Table 4.1): the value for each cell is the number of pairs for which the separations fall within the corresponding two bands.

When these observed counts have been found, the remaining task is to determine a suitable ‘null’ distribution against which the value for each cell can be compared. Though this can be done in several ways, a popular form involves a permutation approach. For some number n_K of iterations, the spatial locations of the incidents are permuted and a new contingency table (of the same form as that for the ob-

served case) is produced for this modified set of incidents. When this process is complete, n_K tables have been produced, each of which represents a realisation of the classification process under the hypothesis that there is no association between the spatial and temporal distributions of the incidents. One of the advantages of the permutation approach is that the spatial and temporal distributions of offending are preserved at each iteration, so that the test is not distorted by clustering present in either of those dimensions independently.

For each cell of the contingency table, the observed value can be compared against the n_K corresponding values in the null-generated tables. The extremity of the observed value, relative to this set of values, is a measure of the extent to which the clustering of the observed incidents departs from what would be expected if their spatial and temporal characteristics were independent. Formally, the statistical significance of the observed value can be estimated by finding the position, R_K , which the observed value would occupy in a rank-ordered list of the null values for the cell and applying the formula

$$p = \frac{n_K - R_K + 1}{n_K + 1} \quad (4.1)$$

for the pseudo p -value (North *et al.*, 2002). The magnitude of any effect can also be estimated either by computing a z -score for the observed value, relative to the distribution, or by finding the ratio of the observed value to the median of the distribution. The former option will be used here.

In this way, the extent of clustering is quantified for every combination of spatial and temporal bands, *e.g.* “in the second week after a burglary has taken place, significantly more incidents tend to occur at a distance between 101 and 200 metres than would be expected on the basis of chance”. Because of this, the results can be used to assess the way in which the magnitude (and presence) of clustering varies over space and time, thereby providing an estimate of the spatial extent of risk elevation.

This test was carried out, with $n_K = 99$ iterations, for the Birmingham burglary data, and the results are shown in Table 4.1. Highly significant clustering is evident at almost all scales examined, and a general (though not monotonic) trend of decreasing influence can be seen at increasing levels of separation. These results are entirely in line with expectation and consistent with those found elsewhere (Johnson *et al.*, 2007).

		Spatial band - upper limit (metres)									
		100	200	300	400	500	600	700	800	900	1000
Temporal band - upper limit (days)	7	9.78	8.35	8.07	8.20	7.21	6.56	5.39	4.39	5.85	3.58
	14	7.79	7.25	6.19	5.56	6.90	6.25	4.35	4.34	5.16	3.63
	21	6.63	5.47	6.10	5.39	5.83	4.76	3.60	3.30	4.30	4.00
	28	5.93	5.07	5.19	5.72	3.78	3.05	4.67	3.80	2.29*	1.62*
	35	5.75	5.71	4.11	4.20	4.67	3.51	3.33	3.49	1.65*	2.24

Table 4.1: Z-scores for a Knox test performed on 26,614 incidents of burglary in Birmingham. Bands are denoted by their upper limit, so that, for example, the band denoted by 200 comprises all values in the range (100,200]. Non-significant values are marked with *; all others are significant with a p -value of 0.01.

4.3.2.2 Network distance

Although the technical definition of space-time clustering - dependence between the spatial and temporal distributions of crime - is a universal one, its precise meaning does, naturally, depend on the way in which space and time are represented. For space in particular, there are well-motivated alternatives to the ‘as-the-crow-flies’ notion of distance commonly used, which might better reflect the role of space in human activity. As has been argued previously in this thesis, there are good reasons to consider that the locations of places in an urban environment might be most appropriately thought of in terms of their position on the street network, particularly when navigability is a concern.

Adapting the Knox test to a network framework is straightforward: all that is required is to alter the distance metric used in pairwise comparisons and to modify

the spatial bands accordingly. Nevertheless, such an adaptation has not previously been used in published work, and so the issue of clustering has not been examined in network terms. Doing so here achieves this, and also provides a more immediate motivation for the work on directionality which follows. Of course, the two measures of distance - network and Euclidean - are highly correlated, and therefore a positive result is to be expected, given the findings of the previous section.

The notion of network distance used here is based on topological separation; *i.e.* the adjacency of streets is the fundamental concern. The distance between any pair of incidents is therefore defined as the number of topological steps between the street segments on which they occur: 0 if the same street, 1 if they occur on neighbouring streets, 2 if there are two degrees of separation, and so on. This is a discrete measurement, and so exact values can be used without the need for spatial bands.

Again, the choice to measure distance using these units is motivated by a combination of technical and theoretical concerns. Keeping in mind the ultimate aim of analysing (near-)repeat targeting, street segments represent a convenient unit for choice-based analysis, and so it is useful to understand clustering in discrete terms. As noted in Section 3.2.1.1, the inclusion of additional granularity (by using fractions of streets, for example) is likely to be counter-productive, since all significant variation in network metrics takes place at the segment level. More practically, though, results stated in terms of degrees of separation represent an appealingly parsimonious, and useful, outcome. A finding expressed in terms of ‘the risk to properties 2 streets away’, for example, provides a contrasting perspective to that of metric distance, and might suggest an easily-comprehensible heuristic for intervention.

The results of this test, using a process identical in all other respects to that of Section 4.3.2.1, are shown in Table 4.2. In this case, the results are found to be highly significant for all levels of separation, as is consistent with the results for the Euclidean case. Particularly strong effects can be seen for the one-week window,

across several degrees of network separation, which provides a useful indication of the extent of such effects. One other observation is that, although there is an apparent decrease in effect magnitude with increasing spatial separation, the relationship is far from monotonic. The fact that this differs from a simple distance-decay relationship provides partial motivation for the study of whether the process is biased towards streets with certain characteristics. Overall, though, the main conclusion of the results is that space-time clustering is clearly evident, and can therefore be well identified in network terms.

		Spatial separation (steps)					
		0	1	2	3	4	5
Temporal band - upper limit (days)	7	9.77	8.65	7.93	7.81	7.57	7.65
	14	7.99	4.71	5.21	6.12	4.65	5.79
	21	6.54	4.14	3.22	5.03	4.98	4.27
	28	5.46	3.27	3.69	5.68	4.20	4.42
	35	6.12	3.38	4.15	5.64	5.25	4.67

Table 4.2: Z -scores for a Knox test performed on 26,614 incidents of burglary in Birmingham, where distance is calculated in terms of degrees of separation between street segments. All values are statistically significant with a p -value of 0.01.

4.3.3 Directionality

That space-time clustering can be demonstrated for networks is a relatively small advance on previously-established results, essentially showing that the phenomenon of clustering can be understood in different terms. The particular value of the approach, however, becomes apparent when considering the logical next step in the analysis and modelling of events of this kind. Given that it has been shown that space-time clustering (and the more specific concept of (near-)repeat victimisation) does occur, it is natural to examine what can be said about *where* such events occur.

Several approaches have been suggested for the prediction of areas at risk from near-repeat victimisation; these include traditional methods based on Kernel Density Estimation (see Chainey & Ratcliffe, 2005) and the ProMap method proposed

by Bowers *et al.* (2004). All of these, however, operate on the principle that any boost effect acts uniformly in all directions; this is unsurprising, given that nothing more can be said on the basis of, for example, the Knox test. These techniques can be refined by adjusting predictions according to opportunity (*i.e.* the distribution of targets), but this is done under an assumption of uniform risk across targets, and takes no account of their relative risk. Again, this may be due to an absence of suitable explanatory factors which could be incorporated into such a calculation.

The properties of the street network, however, do indeed provide a means by which the inherent characteristics of space (in a sense which is relevant to crime) can be quantified. Areas surrounding the location of a crime can be distinguished from each other on the basis of their network properties, and these therefore represent a potential guide for prioritisation. If such a relationship can be found, in that network properties can be seen to influence the directionality of crime spread, then repeat victimisation effects can no longer be considered to be uniform in all directions.

4.3.3.1 The effect of the network

There is indeed good reason, based on criminological theory, to expect that the street network should influence (near-)repeat victimisation. This can be seen by appealing to the key theoretical concepts of the topic: the *flag* and *boost* hypotheses, and the optimal forager principle.

The first of these - the flag hypothesis - is relatively uninteresting in this regard, since its assertion that (near-)repeat victimisation can be explained by fundamental heterogeneity of risk implies that all insights into the issue are simply corollaries of results for the static case. There is therefore little scope to develop hypotheses specific to dynamic effects.

In the case of the boost hypothesis, and related notion of the optimally-foraging offender, however, specific links can be identified. This account of (near-)repeat victimisation is based on the principle that the victimisation of a property actively

increases its risk of victimisation and that of those around it for a short period; that is, it causes them to be more likely or appealing targets. This can be reconciled with the expectation that the surrounding properties are likely to be in the activity space of the initial offender and are therefore more likely to be considered as possible targets for a prospective follow-up offence.

The implication of this argument is that, as with the static case, imbalances in the risk of victimisation can be explained by the tendency of some parts of the street network to feature more often or more prominently in awareness spaces. In the static case, this was estimated by considering the accumulation of all possible journeys through the network, encapsulated by the metric betweenness. Since all journeys were treated equally in these calculations, this was effectively an estimate of the travel activity of the pedestrian population at large: because the outcome of interest was victimisation by *any* offender, this was an appropriate choice.

When estimating awareness spaces in the context of (near-)repeat victimisation, though, there is a crucial difference: the fact of the initial victimisation determines the perspective from which the awareness space should be evaluated. If the prospective follow-up offence is assumed to be the work of the same offender, the task changes from estimating awareness space in general to that of estimating the awareness space of one offender (or group of offenders) in particular. Importantly, one piece of information about this awareness space is known: it includes the location of the initial offence. Of course, this argument relies on simplifying assumptions - the same-offender principle, and a relatively literal and rigid interpretation of the role of awareness space in target choice - but both are crystallisations of concepts for which there is strong support within the literature (*e.g.* Bowers & Johnson, 2004; Bernasco, 2008).

The strategy implied by the above argument can be encapsulated by one question: given that an individual's awareness space includes a certain street, which

other streets are likely to also be in the same awareness space? In terms of travel, this corresponds to asking which streets an individual is likely to travel along, if he or she is known to travel (or recently to have travelled) along a particular street. If this can be estimated, it provides a means of determining, after an initial offence, which surrounding streets the offender is most likely to be aware of and therefore to consider for a follow-up offence. This line of reasoning provides the motivation for the definition of a suitable network metric.

4.3.3.2 Commonality

The analysis of networks is typically performed in terms of the properties or characteristics of individual features (*i.e.* vertices or links). Frequently-used metrics such as degree and betweenness are of this type, and numerous others have been proposed, emphasising various aspects of network structure. There are, however, few methods for the measurement of dyadic relationships; that is, the association between pairs of features. ‘Pairs of features’ is taken here to refer to arbitrary pairs: the presence of a link is, of course, a dyadic relationship between the two terminal vertices, but non-neighbouring features might also be meaningfully compared. Such comparison might, for example, be used to quantify some notion of similarity, or to infer the potential for interaction between two features. The spread of crime risk between streets is, of course, such an interaction.

The case of crime, when considered in detail, exemplifies the need for such a dyadic measure. The argument above suggests that there is value in asking, given a certain link in a network - denoted e , for concreteness - which other links also tend to feature in journeys which use e . Simply considering the individual properties of these links is insufficient, and might be misleading. To take the example of betweenness, two other links e' and e'' could have the same betweenness value but relate very differently to e : e' might be an immediate neighbour, and e'' might be in an entirely different part of the network. Immediate neighbours are not all equal, either: e may be connected to a *cul-de-sac* at one end and a main road at the other. These problems can only be addressed by considering the centrality of links from the perspective of e .

To this end, a new measure of the *commonality* of two links is introduced. Its definition can best be understood as a derivative of betweenness: where the betweenness of a link e counts the number of shortest paths which include it, the commonality of e with another link e' calculates the proportion of these paths which also feature e' . In essence, then, it measures the co-occurrence in paths of any pair of links on the network, normalised by the overall occurrence of one of those links. Expressed another way, it measures the proportion of shortest paths passing through e which also pass through e' .

Commonality can be defined formally using many of the same terms used to specify betweenness in equation (3.1). For generic vertices v and w , and for any pair of links e_i and e_j , σ_{vw} is defined as the number of shortest paths between v and w , and $\sigma_{vw}(e_i)$ is the number of those shortest paths **which pass through** e_i . If a further definition is added for $\sigma_{vw}(e_i, e_j)$ as the number of shortest paths between v and w which feature **both** e_i and e_j , then the commonality of e_j relative to e_i , denoted C_{ij} , is defined as

$$C_{ij} = \frac{\sum_{v \sim w} \frac{\sigma_{vw}(e_i, e_j)}{\sigma_{vw}}}{\sum_{v \sim w} \frac{\sigma_{vw}(e_i)}{\sigma_{vw}}}, \quad (4.2)$$

As before, \sim represents the relation ‘there exists a path between v and w ’.

A number of basic properties of commonality are immediately apparent. The denominator of C_{ij} is simply the betweenness of link e_i , and so C_{ij} can be thought of as the proportion of the journeys that contribute to the betweenness of e_i which also incorporate e_j . As such, values lie in the range $[0, 1]$, and the upper limit of 1 is realised at least once for every i : trivially, $\sigma_{vw}(e_i, e_i) = \sigma_{vw}(e_i)$ and so $C_{ii} = 1$ for all values of i .

It is also important to note that the definition of C_{ij} is not symmetric in i and j : it refers specifically to the commonality of e_j relative to e_i . Such asymmetry is

evident in the real-world situations to which the measure responds: to cite another example from street networks, where a busy street and quiet street are adjacent, the dependence of the quiet road’s traffic on the busy road is likely to be greater than that in the opposite direction. This point will be illustrated using real-world examples in the following section.

Commonality has one further advantageous feature, which concerns the fact that a distance-decay relationship is implicit within its definition. Models and theories of interactions in space typically incorporate some notion of distance-decay, whereby the strength of a relationship varies inversely with spatial separation. This is evident in journey-to-crime data for many criminal phenomena (Wiles & Costello, 2000), and is a ubiquitous feature of near-repeat victimisation (see Tables 4.1 and 4.2, for example). Often this must be accounted for explicitly (by including a ‘distance’ variable, for example), but consideration of commonality reveals that it is an effect which arises spontaneously. For a given link e_i , values of C_{ij} will tend to be lower for links e_j which are further removed: moving further away from e_i , the plurality of possible routes increases with every junction, so that the load from e_i will become more dispersed.

The effect can be seen most clearly by considering extreme cases. All but one of the journeys which use e_i will feature at least one of its immediate neighbours (the exception is the trivial journey from one end of e_i to the other) and so C_{ij} will tend to be high for such an e_j . On the other hand, if e_j is a distant *cul-de-sac* then it will feature in a very small proportion of e_i ’s journeys and give rise to a correspondingly low C_{ij} . This property is an important feature of commonality. Although the ability to compare links in a manner which accounts for their relative location in the network appears modest, it represents a significant improvement on approaches which rely exclusively on immediate adjacency.

The final point to note is that, analogously with betweenness, the definition of

commonality can be refined in order to consider only trips whose length is less than some maximum radius r . This is done in the natural way - by modifying the definition of \sim in (4.2) to incorporate only pairs of vertices which lie within r of each other - and the resulting measurement is denoted $C_{ij}^{(r)}$. As with betweenness, r can be measured in either metric or topological units, and the distinction will be made whenever it is used. Again, restricting the maximum path length can be interpreted as representing more localised phenomena (pedestrian travel, for example) and as a means of ameliorating edge effects.

4.3.3.3 Real-world examples

As with many network metrics, the meaning and value of commonality can perhaps best be understood by considering how it is applied to a real-world example. In this section, examples from the street network of Birmingham will be used to illustrate its main properties. Figure 4.1 illustrates, using one section of the street network of Birmingham, the extent to which the centralities of street segments vary depending on the perspective from which they are viewed. In order to establish a basis, 4.1a shows the betweenness of the links: this can be considered to be a ‘global’ perspective, and represents the extent of understanding which can be gained by considering the properties of links individually. The section shown includes an intersection between two highly-between roads (which appear to be arterial routes flowing north-south and east-west) and a number of more isolated side-roads with low betweenness values.

Figures 4.1b to 4.1e show commonality from the perspective of four particular links - e_1, e_2, e_3 and e_4 - and demonstrate that the true nature of traffic flow is somewhat more nuanced. In particular, comparison of 4.1b and 4.1c reveals the disparity in flow patterns at the crossroads: e_1 and e_2 play identical roles at the junction, and have very similar betweenness, yet there is a clear difference in the routes which use them. On the basis of commonality, it can be seen that the majority of journeys which use e_1 also feature other links on the main north-south route; that is, it is journeys on this route that are responsible for the majority of the betweenness of e_1 .

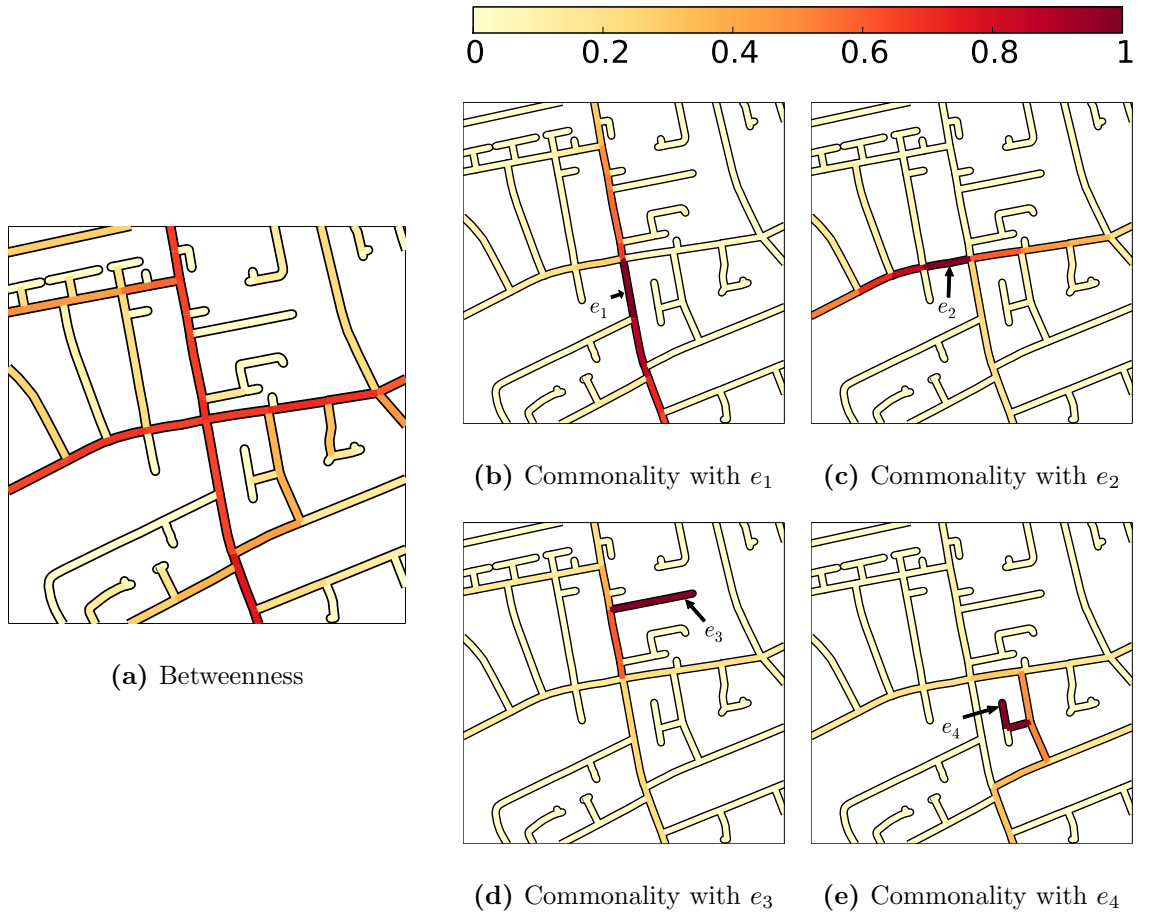


Figure 4.1: Betweenness and commonality for a small section of the street network of Birmingham, using measures based on a limited radius of 2,000 metres. Panel a) shows links coloured according to betweenness, the notion of centrality from which commonality is derived. In panels b) to e), links are coloured according to their commonality relative to a specific choice of link - respectively e_1 , e_2 , e_3 and e_4 - showing the degree to which they feature in the same set of journeys. Colours therefore reflect values of $C_{ij}^{(r)}$ for specific choices of i .

Relatively few of the journeys which use e_1 involve a turn at the central junction, and the same is true of e_2 . The notion of commonality is necessary to reveal this: from the perspective of e_1 , all three adjacent links at the central junction have the same betweenness and are therefore indistinguishable on that basis.

Figures 4.1d and 4.1e show further examples of the added insight afforded by commonality. Link e_3 is a low-betweenness segment which is immediately adjacent to a highly-between route; in addition, because it is a *cul-de-sac*, all journeys which use it must also include one of the two segments to which it is adjacent. Commonality reveals, however, that the load is not evenly distributed between the two, as it can

be seen that the more southerly of the two features in a greater proportion of trips. The most obvious reason for this is that the segment connects e_3 to the main local crossroad, and so will be used as journeys are ‘funnelled’ towards the main roads that meet there. This demonstrates another sense in which the commonality of two links is not simply a function of the distance between them.

Figure 4.1e shows a similar, though more extreme, case. Link e_4 is, again, a *cul-de-sac*, but is not immediately adjacent to a high-betweenness link. The commonality of one neighbour is only negligibly less than 1, since almost all journeys which use e_4 originate/terminate somewhere in the main part of the network and must therefore incorporate that link. The other neighbour of e_4 has only very small commonality, since only one shortest path uses both links. This can be reconciled with human travel patterns: a pedestrian who traverses e_4 is virtually certain to have also used the high-commonality neighbour, whereas only one particular circumstance would involve the use of the other neighbour.

Each of the other situations can be translated straightforwardly into the language of pedestrian movement and awareness spaces. If a pedestrian is known to use e_1 , for example, it can be said with high probability that his or her awareness space also includes other parts of the main north-south route. Other segments, however, are only used in a small number of circumstances, and are therefore less likely to be in the awareness space of a randomly-selected user of e_1 .

4.3.3.4 Analysis via discrete choice

The commonality metric introduced above provides a means of estimating the extent to which network links tend to co-occur in travel patterns and, by extension, in the awareness spaces of pedestrians. According to the theoretical argument outlined in Section 4.3.3.1, this should correspond, to some extent, to the elevation in risk experienced by nearby locations in the aftermath of an initial victimisation. There are a number of methods by which this hypothesis can be tested.

The aspect of (near-)repeat victimisation over which commonality is expected to exert influence is the location at which the follow-up incident occurs (in practical terms, the street segment on which it occurs). It is, therefore, natural to take the occurrence of such a (near-)repeat pair as the starting point for analysis; that is, to examine whether, among all such pairs of incidents, there is a trend in the location of the second incident. Implicit in this approach is that the occurrence of the secondary offence is a *fait accompli*; that is, that only the location of the second offence is in question, rather than its occurrence. This echoes the use of discrete choice models elsewhere in criminology, where target choice is treated independently of the initial decision to offend (*e.g.* Bernasco, 2009). It should be noted that, at this stage, the definition of a (near-)repeat is generic: it refers only to a pair of incidents which occur within D streets and T days of each other. For consistency, the two incidents will be referred to as ‘initial’ and ‘secondary’.

When these are identified in data, all pairs which meet the criteria are included; this allows the possibility, for example, that the same crime may be the secondary incident in more than one pair. Although this may be considered undesirable, the absence of a method by which crimes can definitively be linked means that some compromise must be made in this respect. Two arguments can be made in favour of using all pairs. On one hand, where an incident appears multiple times, it may be the case that all linked pairs in which it is involved are valid (in the sense of a common offender) and so all are the result of a common targeting process. In a more practical sense, this corresponds to what would be available in a predictive scenario: given an incident, it is possible that it may trigger multiple near-repeats in different locations.

Working with pairs of incidents, defined as above, as the fundamental unit, there are a number of ways in which the locations of the secondary incidents could be analysed. One possibility would be to examine the characteristics of the places at which they occur, either in terms of their network properties or otherwise. Although it may be possible, by this method, to identify features which tend to be shared by

secondary targets, it overlooks a crucial aspect of the situation; namely, that there are a number of other possible locations which were *not* chosen. If an initial incident is to be followed by a secondary victimisation within T days, there are a number of distinct ways in which this can be realised: an offence on any of the segments within D steps of the initial victimisation would satisfy the definition of ‘close pair’. The fact that, in reality, the secondary incident happens on one segment in particular implies that the segment has been ‘chosen’, in some sense, from the set of possible locations. By examining the characteristics of the chosen segment relative to those not chosen, it may be possible to infer the influence of those characteristics on the choice.

This line of reasoning has been presented in previous criminological research, notably by Bernasco & Nieuwbeerta (2005) and Clare *et al.* (2009). In those papers, both of which concern location choice for the crime of burglary, the authors suggest that the characteristics of target locations should most appropriately be thought of not as dependent variables, but as independent variables which influence the outcome of a decision process. The decision process in those cases concerns target choice relative to offenders’ home locations - the focus of the work is target choice in general, rather than the specific case of (near-)repeat victimisation - however, it is easily seen that this is conceptually identical to the present setting, with the role of offender homes assumed by the locations of initial incidents. A similar framework ought, therefore, to be applicable here.

The framework in question is discrete choice modelling, which was proposed by McFadden (1984) as an econometric technique and has subsequently been applied in a wide array of contexts. The approach applies to situations in which a single selection is to be made from a set of discrete options, by a rational actor. The rationality of the actor represents the assumption that the choice is made on the basis of reasoned comparison of the properties of the options, *i.e.* by engaging in some form of cost/benefit calculation and choosing the option perceived as most beneficial. The approach is therefore predicated on the idea that, for any chooser, the

perceived benefit of any option can be quantified as a *utility*, for which the option with the highest value is then chosen. The form of the utility calculation - that is, the factors which it incorporates and their relative weightings - therefore encodes the essence of the decision process, and it is that which analysis seeks to establish.

A clear link can be drawn between the notion of rationality implicit in the definition of discrete choice models and the rational choice perspective within criminology (Cornish & Clarke, 1986). That perspective, which is a common foundational theme throughout this thesis, suggests that all criminal events can be conceived as the outcome of a decision process (of varying levels of sophistication). The appropriateness of the discrete choice framework can be justified by appealing to this, which, in spatial terms, corresponds to evaluation of the trade-off of costs and potential rewards at each location. As applied to crime, the rational choice perspective typically represents offender decisions as an imperfect process, due to bounded rationality on the part of offenders, their incomplete access to information, and the observer's inability to account fully for the utility calculations which occur. In discrete choice models, these are accounted for by the fact that utility is taken to have a random component, representing unobserved factors, which varies between choosers.

It may be noted, at this stage, that the decision process hypothesised here is somewhat more abstract than the traditional notion of choice: that of a situation in which a chooser is presented with a set of overt options and selects from within this. In the context of target choice, the prospective offender is not necessarily assumed to be equally aware of all possible options when the decision is made. Indeed, the likelihood of awareness is precisely the kind of segment-level characteristic whose influence is to be evaluated (and whose importance is predicted by theory). In fact, the choice could be framed in these terms: if victimisation is taken to be directly related to awareness, the choice might actually be defined as 'the choice of which targets to become aware of during routine activities'.

However it is conceptualised, it is clear that the location at which a secondary offence takes place is the realisation of a process for which several outcomes were possible *a priori*, and that rational behaviour plays some part in its selection. The discrete choice framework is applied on this basis. Several precedents for this can be found within the literature: as well as the aforementioned work on burglary, it has also been used in studies of street robbery (Bernasco, 2009; Bernasco & Block, 2009) and civil unrest (Baudains *et al.*, 2013b). Notably, the issue of target awareness is germane in these cases also, since several of them feature choice sets of such size that active consideration of all options is implausible. Indeed, one of the findings which is common to all these studies - that distance acts as a significant disincentive - can be interpreted as relating to the deterioration of awareness with distance.

4.3.3.5 The conditional logit model

The class of ‘discrete choice models’ encompasses a wide range of particular variants, which correspond to differences in the structure of the choice set and in the types of variation accounted for by variables. A full review of these, and the means by which they are estimated, is provided by Train (2003). The form used here is the *conditional logit*, which is applicable in situations for which the choice set contains 3 or more items and where the terms of the utility function are allowed to vary over alternatives. It is this model which is used in the studies cited in the previous section, and its structure will now be introduced in the context of the present scenario.

The elements of the model correspond to the individual instances of choice which it describes; in the present analysis, each of these is a (near-)repeat pair of incidents. These are indexed by k , so that choice k represents the target selection for the secondary incident in the k th pair. This could also be regarded, perhaps more intuitively, as an indexing of choosers: chooser k is the offender who commits the secondary offence in pair k .

The index l is then taken to indicate a member of the set of alternatives available to offender k . For the (near-)repeat scenario, this set comprises all street segments on

which the secondary incident could have taken place while satisfying the definition of (near-)repeat, *i.e.* all segments within D steps of the initially-victimised segment. The choice set is not, therefore, constant across all offenders; rather, it depends on the location of the initial incident in each case.

In order to apply a principle of utility-maximisation, it is necessary to associate a utility with every alternative with which each offender is presented. Accordingly, for all possible combinations of offender k and segment l , the utility of k choosing l is given by

$$U_{kl} = V_{kl} + \epsilon_{kl}, \quad (4.3)$$

where V_{kl} is a *representative utility* and ϵ_{kl} is an *error* which captures the remainder of U_{kl} , *i.e.* that which is not attributable to V_{kl} . The role of V_{kl} is to represent the component of utility which is due to factors observed by the researcher. Though the values of those factors are known, the form of V_{kl} and the values of its constituent parameters are not, and it is these which are to be estimated statistically. These values are taken to be constant throughout the model, and one of the roles of ϵ_{kl} is therefore to account for variation which is idiosyncratic to each offender. More generally, ϵ_{kl} accounts for the effect of all unobserved factors upon the utility U_{kl} , and is modelled as a random variable.

In this model, V_{kl} is taken to be a linear combination of the observed attributes of alternative l , as perceived by offender k . If there are H such properties, with X_{hkl} taken to denote the value of attribute h perceived by offender k for his/her option l , then V_{kl} is given by

$$V_{kl} = \sum_{h=1}^H \beta_h X_{hkl}, \quad (4.4)$$

where the β_h are the coefficients associated with each attribute h . It is worth highlighting at this stage that the X_{hkl} are defined explicitly in terms of how they are perceived by the offender in question; *i.e.* their values are allowed to vary with k . When one of the attributes represents distance, for example, it is calculated in these terms, as the distance between alternative l and the location at which k is making

the choice. This notion of relative measurement is especially relevant here since it corresponds directly to the underlying principle of the commonality metric.

It is clear from equation (4.4) that it is the coefficients β_h which encode the extent of the roles played by each of the H attributes in the evaluation of alternatives by offenders. A value of $\beta_h = 0$, for example, indicates that attribute h has no effect; otherwise, the magnitude of a non-zero β_h quantifies the strength of its influence. Establishing the coefficients of V_{kl} is, therefore, the objective of analysis.

In this model, it is assumed that the error terms ϵ_{kl} are independently and identically distributed according to an extreme-value type one distribution (Gumbel distribution). This being the case, it is possible to write down an expression for the probability of a given offender k choosing an available alternative l . If Y_k is taken to denote the choice of offender k and L_k is the set of alternatives available to k , then it can be shown (see Train, 2003) that

$$\begin{aligned} \mathbf{P}(Y_k = l) &= \frac{\exp(V_{kl})}{\sum_{m \in L_k} \exp(V_{km})} \\ &= \frac{\exp(\beta_1 X_{1kl} + \beta_2 X_{2kl} + \dots + \beta_H X_{Hkl})}{\sum_{m \in L_k} \exp(\beta_1 X_{1km} + \beta_2 X_{2km} + \dots + \beta_H X_{Hkm})}. \end{aligned} \quad (4.5)$$

This can be understood intuitively: after all utilities are exponentiated, the probability of choosing l is the value for l as a proportion of the cumulative total across all alternatives available to k . It is also straightforward to interpret the partial coefficients β_h as determining the effect on this probability of changes in the attribute values X_{hkl} .

Equations (4.3), (4.4) and (4.5) together define the substance of the conditional logit model, and provide the means by which the coefficients β_h can be estimated on the basis of data (using a randomly-distributed ϵ_{kl} , as indicated above). This can be done via maximum likelihood estimation, and this is implemented in a number of software packages. Throughout the remainder of the work, estimates are generated using the program Biogeme (Bierlaire, 2003).

4.3.3.6 Discrete choice results for Birmingham

Having motivated the use of a discrete choice framework as a means of testing the hypotheses of this chapter, it remains to examine the results obtained statistically from the Birmingham burglary data. So far, the conditional logit model has only been introduced in general terms, and many details remain to be specified, such as the variables included and the data points to which it is applied. Both of these will be discussed before the numerical results are presented.

Model structure The specific form of the conditional logit model is determined by the equation for V_{kl} , the representative utility of any alternative l , as introduced in (4.4). A linear form is used here, and therefore all that is required is to establish which variables are to be included.

Given that the primary focus of this work concerns the role of network structure in risk elevation, variables which represent network metrics are present at the core of the model. In order for any results related to these to be meaningful, though, it is necessary to control for other factors likely to have a material effect on crime occurrence by including additional variables. In the model for static risk presented in Chapter 3, a number of demographic factors (*e.g.* unemployment rate) were included for this reason, and found to be significant. In the present situation, though, these factors are not as relevant (or technically expedient): the spatial units which form the choice set here are contained in a very small area and there is therefore little (or no) variation in such factors across choices. Since they are of little discriminatory value, they are omitted in the interest of parsimony.

One factor which does vary significantly from street to street is opportunity, and this can be controlled for by including in the model the number of potential burglary targets on each segment in the choice set. For any alternative l , therefore, the count of residential address points is included as an independent variable. This is an intrinsic property of the segment and therefore does not vary from offender to offender.

Network effects are incorporated via the commonality metric introduced in Section 4.3.3.2. In this case, the value of the variable depends on both k , the identity of the offender (or, rather, the location of the initial offence), and the possible location l being considered. Since the notion of commonality is based on the idea of measuring the centrality of one segment relative to another, it applies straightforwardly to this case by considering the value of the candidate segment from the perspective of the initially-victimised segment. In terms of notation, commonality is defined in terms of pairs of segments, so that C_{ij} is defined as the commonality of segment e_j relative to e_i . If attribute h represents commonality, therefore, X_{hkl} is equal to C_{ij} , where e_i is the segment on which the k th initial offence took place and e_j is the segment represented by alternative l .

The version of commonality which is used in the statistical models is based on a radial limit of 2,000 metres. Experiments were performed with different radii; however, the general findings varied little from case to case and therefore results for these versions of the models are omitted for brevity. Differences in the precise definition of commonality are not the primary interest here - this work simply addresses the question of whether the general principle is of explanatory value - but will be explored in later work. The particular value of 2,000 metres was chosen because it is close to the upper limit of what was possible with the computational resources available: the calculation of commonality is computationally expensive and scales super-linearly with radius. The choice of 2,000 metres does, however, lie within the range of values for which betweenness was found to be significant in Chapter 3.

One question which does arise at this stage is whether to include an independent variable representing inter-segment distance in the model. Given the theoretical importance of the distance decay principle (for near-repeat victimisation in particular), such a variable would seem to be an essential component of any model. However, as pointed out in Section 4.3.3.2, distance decay is implicit, to an extent, in the defini-

tion of commonality. For a given segment, i , the commonality of another segment j is inversely related to the distance between i and j , and therefore such effects may already be accounted for.

While this relationship is true in general, though, the exact scaling of commonality with distance varies from case to case. To investigate the extent to which commonality does capture both effects, two different models are examined: one in which commonality and distance are both included as independent variables, and one in which distance is omitted. The measure of distance used in this case is topological, *i.e.* the number of degrees of separation between the two segments. A summary of the models is given in Table 4.3.

Model number	Independent variables
1	Address count, commonality
2	Address count, commonality, distance

Table 4.3: Independent variables included in the two conditional logit models considered.

Data points It is also necessary to specify the data points which were used in order to estimate the model statistically. Each of these points is a ‘close pair’ of incidents, which is then used to define a choice problem by considering the possible locations at which the secondary incident could have taken place. The issue therefore concerns the precise definition of ‘close pair’ and the set of these which were considered for each model.

The temporal aspect of the close pair definition was kept constant through the analysis, and a close temporal pair was taken to be one for which the incidents were separated by between 1 and 7 days (inclusive). The definition was kept constant so as not to complicate the analysis by introducing further variation, though other intervals may be considered in further work. The particular choice of 7 day interval was made due to its prevalence in related work on the subject; in addition, it is a

level at which near-repeat effects are observed universally. Same-day offences were omitted in order that the temporal precedence of any pair of incidents (which is ‘initial’ and which ‘secondary’) was known definitively.

As regards the spatial aspect of the close pair definition, several options were explored. As in Section 4.3.2.2, the definition is made in terms of ‘degrees of separation’, so that a value of 0 corresponds to the same street, 1 to an immediate neighbour, and so on. In the first case, a close spatial pair was considered to be one for which the incidents occur within 3 street segments of each other, but this was then modified in order to discount same-segment offences (so that the separations considered were 1-3 inclusive).

With a specific definition of close pair established, it was straightforward to construct the dataset for choice analysis. For each close pair, the choice set consisted of all street segments which lie within the designated spatial radius of the initially-victimised segment and have at least one address point (*i.e.* all viable targets within the spatial bandwidth). Relevant metrics were calculated for all of these and, together with the observed outcome (the true victimised segment), these comprised the data point.

Unfortunately, computational restrictions meant that it was impossible to carry out the analysis for the city of Birmingham as a whole. As noted above, the resources required to calculate commonality scale super-linearly with network size and maximum trip distance, and it was infeasible to carry this out for the whole network of Birmingham with the resources available. Instead, the analysis was carried out for smaller contiguous sections of the network. The upper limit of feasibility corresponds loosely to the size of Birmingham’s Local Policing Units (LPUs; administrative units based on command structure) and these were therefore taken to be appropriate sub-units. Results are shown for Birmingham East, North and South LPUs, which give rise to contrasting findings. The other LPU, Birmingham West &

Central, includes the city centre of Birmingham, which is dominated by retail and has an intricate traffic system including many motorways; analysis is likely to be least reliable in such an area and it is therefore omitted for brevity.

Numerical results In the first case, results were estimated using the set of close pairs identified as occurring within 1-7 days and 0-3 street segments of each other. As seen in the earlier Table 4.2, highly significant clustering is present at this bandwidth in Birmingham as a whole. The results of the conditional logit models in this case are shown in Table 4.4.

	Birmingham East		Birmingham North		Birmingham South	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Independent variables						
Address count ($\times 100$)	1.97* (25.05)	1.91* (24.27)	2.18* (19.35)	2.12* (18.51)	1.93* (41.03)	1.89* (39.46)
Commonality	0.734* (9.84)	0.353* (3.76)	0.728* (7.42)	0.440* (3.43)	0.559* (8.38)	0.189* (2.26)
Distance (degrees of separation)		-0.198* (-5.78)		-0.15* (-3.18)		-0.191* (-6.16)
Goodness-of-fit						
Number of pairs	1655	1655	903	903	2001	2001
Initial log-likelihood	-5149.22	-5149.22	-2727.17	-2727.17	-6149.08	-6149.08
Final log-likelihood	-4664.09	-4647.33	-2521.71	-2516.29	-5358.637	-5339.84
Likelihood ratio	970.26	1003.76	410.92	421.77	1580.88	1618.46

Table 4.4: Results for conditional logit models where close pairs are defined as those occurring up to 3 street segments apart, including same-segment offences. The values shown are estimates for the β_h coefficients for each variable, and the values in brackets give the robust t -statistic in each case. An asterisk * denotes a value which is significant at $p < 0.01$ level.

Across both models and all three areas of Birmingham, commonality is seen to be a highly significant predictor of location choice. This is consistent with the hypothesised relationship, and suggests that commonality does indeed have potential

as a means of predicting risk elevation in the aftermath of a burglary. In all cases, the predictive power of the model is increased by the addition of distance - the log-likelihood of Model 2 is greater than that of Model 1 for all three areas - suggesting that distance-decay is both a significant factor and also not one which is captured fully by commonality. Nevertheless, commonality remains highly significant even when distance is included. It is notable, however, that results do vary across areas: the coefficient associated with commonality is noticeably lower for Birmingham South than in the other cases, for example.

The results of Table 4.4 are encouraging in terms of their support for the use of commonality. One concern does arise, however, when considering the appropriateness of the spatial bandwidth used in the analysis. This arises from the fact that ‘same segment’ values are something of a special case for commonality: as remarked in Section 4.3.3.2, C_{ii} is trivially equal to 1 for all values of i . When set against the fact that same-segment offending is disproportionately common in the data (partly as an artefact of distance-decay, and partly due to the specific phenomenon of repeat victimisation), this becomes somewhat problematic. To be clear, it may be that the observed significance of commonality is simply because it has an anomalously high value for segments which are known to be disproportionately victimised anyway.

For this reason, further analysis was carried out in which same-segment offending was omitted. The results in Table 4.5, therefore, were generated for pairs of events occurring within 1-7 days and 1-3 street segments (inclusive in both cases) of each other. This reduces the sample size but ensures that the special case of commonality is avoided.

As anticipated, several relationships become somewhat weaker when this modification is made. The effect of commonality remains significant in both East and North, but at a generally lower level than that seen previously and with notably reduced magnitude. In addition, the effect is no longer significant in Birmingham

	Birmingham East		Birmingham North		Birmingham South	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Independent variables						
Address count ($\times 100$)	1.99* (22.93)	1.97* (22.74)	2.23* (17.88)	2.22* (17.71)	1.91* (38.42)	1.91* (38.13)
Commonality	0.369* (3.48)	0.226** (1.97)	0.391* (2.84)	0.292** (1.92)	0.067 (0.62)	-0.006 (-0.05)
Distance (degrees of separation)		-0.123* (-3.15)		-0.09** (-1.66)		-0.067** (-1.84)
Goodness-of-fit						
Number of pairs	1414	1414	769	769	1650	1650
Initial log-likelihood	-4376.98	-4376.98	-2291.71	-2291.71	-5048.12	-5048.12
Final log-likelihood	-4018.40	-4013.36	-2145.94	-2144.47	-4441.12	-4439.43
Likelihood ratio	717.17	727.25	291.53	294.49	1214.00	1217.39

Table 4.5: Results for conditional logit models where close pairs are defined as those occurring between 1 and 3 street segments apart, *i.e.* not including same-segment offences. The values shown are estimates for the β_h coefficients for each variable, and the values in brackets give the robust t -statistic in each case. An asterisk * denotes a value which is significant at $p < 0.01$ level, whereas a double asterisk ** represents $0.01 < p < 0.05$.

South (where the weakest relationship was found in the initial analysis). The general reduction in effect suggests that the same-segment issue is responsible for a substantial portion of the effects seen in the initial analysis, and indeed accounts entirely for this effect in South. The fact that significant results remain in East and North, though, implies that commonality retains explanatory value in these cases. The loss of magnitude for the distance effect may also be ascribed to the same-segment effect; it is unlikely that the decay relationship identified in the initial analysis was a linear one.

There is no immediately-apparent reason for the discrepancy between areas in the effect of commonality, though several possibilities can be identified. The observed disparity implies that either the character of the network (as measured by common-

ality) is different in Birmingham South, or that the nature of offending is different, or both. The first of these could arise either because commonality is not an effective discriminant in that case (loads are evenly balanced) or because those segments which it identifies are unfavourable for crime for other reasons. On the other hand, the difference may be attributable to different modes of offending: commonality relates specifically to the boost account of repeat victimisation, and if this is responsible to a lesser extent in Birmingham South then that may also explain the difference.

	Birmingham East		Birmingham North		Birmingham South	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Independent variables						
Address count ($\times 100$)	2.32* (16.40)	2.31* (16.32)	2.23* (11.31)	2.22* (11.22)	1.90* (22.38)	1.91* (22.46)
Commonality	0.319* (2.43)	0.245** (1.76)	0.310** (1.85)	0.226 (1.25)	0.070 (0.55)	0.0157 (0.12)
Distance (degrees of separation)		-0.125 (-1.51)		-0.159 (-1.43)		-0.105 (-1.31)
Goodness-of-fit						
Number of pairs	794	794	426	426	886	886
Initial log-likelihood	-1921.28	-1921.28	-979.54	-979.54	-2105.12	-2105.12
Final log-likelihood	-1713.80	-1712.64	-916.44	-915.41	-1848.24	-1847.35
Likelihood ratio	414.98	417.29	126.20	128.26	513.75	515.53

Table 4.6: Results for conditional logit models where close pairs are defined as those occurring between 1 and 3 street segments apart, *i.e.* not including same-segment offences. The values shown are estimates for the β_h coefficients for each variable, and the values in brackets give the robust t -statistic in each case. An asterisk * denotes a value which is significant at $p < 0.01$ level, whereas a double asterisk ** represents $0.01 < p < 0.05$.

One final refinement of the analysis concerns a further modification to the definition of a spatial close pair in which only incidents either 1 or 2 segments apart are considered (*i.e.* cases 3 segments apart are omitted). This restricts the analy-

sis to more local effects, and estimated values are given in 4.6. Effect size is seen to decrease further in these cases, although the significance of commonality is still present in three of the models. At the same time, distance is no longer significant, for a p -value of 0.05, in any of the three areas. This suggests a complex interaction between the two factors: the non-significance of distance here implies that distance-decay effects between distances 1 and 2 are accounted for by commonality (where it is significant), in contrast with the situation when distances of 3 are included. One possible explanation is that the ability of commonality to capture distance-decay is lost at higher distances. This accords with the tendency of commonality to decay fairly rapidly with distance, as loads are spread across numerous routes. The effectiveness of commonality as a predictor might therefore be limited in its spatial range.

4.4 Discussion

The purpose of this chapter was to extend the analysis of the influence of the street network on crime to consider dynamic, rather than static, effects. Space-time clustering, via the phenomena of repeat and near-repeat victimisation, represents a crucial issue in the modelling of crime over short time-scales, and is a feature which invites practical intervention. Understanding how related series of incidents occur is therefore of great value for crime prediction, and has the potential to influence mathematical models of crime.

There is good reason to expect that the street network should influence aspects of (near-)repeat victimisation - specifically, the spatial distribution of secondary offences - since it plays a fundamental role in shaping the awareness spaces of potential offenders. However, the fact that (near-)repeat pairs are likely to be the work of the same offender represents a crucial difference between this analysis and the evaluation of long-term aggregated risk in Chapter 3. In particular, it implies that the notion of awareness space most appropriate to the analysis of secondary target choice is one which is calculated relative to the location of the initial offence.

This motivated the proposal of a novel network metric, commonality, which quantifies the relationship between any pair of network features in terms of their roles in travel through the network. Since it is based on a similar shortest-path principle, it can be regarded as a form of ‘relative betweenness’. It measures the extent to which features co-occur on paths through the network, and therefore can be interpreted as quantifying the likelihood that two streets will be common to the same awareness spaces of individuals travelling through the network.

Preliminary analysis of commonality for the street network of Birmingham showed that it corresponds well to intuitive notions of centrality and of the strength of relationship between elements of the street network. Moreover, the additional insight afforded by considering centrality in relative terms was demonstrated by comparison with betweenness. Whereas betweenness gives no indication as to the nature of the traffic flowing through a link, commonality reveals how links are used in longer journeys through the network. It can therefore be used to differentiate between network features which would be equivalent in terms of betweenness, a non-relative measure.

The relationship between commonality and secondary target choice was then analysed statistically, using a discrete choice model. In general, commonality was found to be a statistically significant predictor of the locations of follow-up offences; specifically, secondary incidents tended to occur on street segments of higher commonality with the initially-victimised segment. This relationship was not, however, universal: its significance and strength was dependent on the area of Birmingham considered and the particular definition of ‘close pair’ used.

The substantial contributions of the chapter are three-fold: the demonstration of space-time clustering in network terms, the conception of the commonality metric, and the statistical analysis of near-repeat data. Several matters arise from these, both in terms of criminology and the study of networks in general.

4.4.1 The relationship between commonality and (near-)repeat offending

The results of the statistical model can be interpreted in two ways, depending on perspective. If the principles of environmental criminology (such as pattern theory) are assumed to hold, the results suggest that the features of the street network which should influence target choice in this context are encoded successfully by commonality. Conversely, if commonality is taken to be a valid quantification of awareness space, the findings offer support for pattern theory. In either case, though, the existence of a relationship suggests that it can be used to inform the prediction of crime locations.

Nevertheless, the sensitivity of the relationship to variation in the setup of the statistical model implies that the results should be treated with caution. For example, the fact that the exclusion of same-segment offence pairs causes a reduction in the explanatory value of commonality (or, for one area, the elimination of the relationship) highlights one unfortunate property of commonality. The commonality of one segment with itself - in matrix terms, each diagonal value - is equal to one, which is anomalously high in comparison with off-diagonal values. This raises the possibility that the success of a commonality-based model can simply be ascribed to its identification of same-segment offending (which is the dominant form of near-repeat victimisation).

This property of commonality is an inconvenient one, and it is impossible to know, within the scope of the present analysis, the extent to which it compromises results when same-segment victimisation is included. One possible remedy, which may be investigated in future work, would be to allow the diagonal values of commonality to be varied as a tunable parameter, which could be adjusted to reflect the true magnitude of the same-segment effect. While this may be analytically expedient, though, it would compromise the simplicity of the metric somewhat, thereby also

limiting its interpretation. An alternative would be to introduce an additional stage in the choice process, representing the decision of whether to offend on the same segment or elsewhere.

A further potential limitation concerns the fact that the relationship between commonality and offending varies between areas. Possible explanations for this may lie in the variability between areas of both street network structure and offender characteristics. In the first case, it may be that certain street networks are simply not amenable to analysis of this type, perhaps due to the presence of idiosyncratic features (although none are apparent in the cases considered here). On the other hand, the hypothesis that offences should occur on streets of higher commonality applies only to incidents which can be ascribed to the boost hypothesis (and, more strictly, to those which are the work of the same offender). If this is violated, then it is to be expected that no relationship with commonality should exist.

This highlights a more general point: that trends in crime patterns can be contingent on aspects of offender behaviour which are not observable *a priori*. In this case, it is unknown whether secondary offences are predominantly committed by the same offender, and yet the viability of the model depends on this being the case. Indeed, this could be cast in terms of an affliction of complex systems more generally: if an emergent phenomenon (space-time clustering) has multiple causes which are sufficient, but not necessary (flag *vs.* boost), models based on only one of these can be inadequate.

In the context of the present work, there are two routes forward. The model can, of course, be tested on data for which the premise is known to be true (*i.e.* offences cleared to the same offender). While this could establish its validity in those circumstances, though, it is of little help as regards practical implementation. For that, it is required to be known: a) to what extent the ‘boost’ hypothesis applies in a given situation, and b) how to make predictions on the basis of both hypotheses together.

Both of these will be considered in future work.

The issue of variation between areas was seen again in a further stage of analysis, in which only incidents separated by 1 or 2 segments were included. This can be reasoned in much the same way as above; however, the fact that distance is not a significant factor in these cases is notable. One possible explanation for this is that the decay in risk between 1 and 2 streets is accounted for by the distance-decay property of commonality (which remains significant in East and North). The possibility that commonality might have an ‘effective range’ may need to be considered in practical implementations.

Overall, the finding that commonality is a significant factor in the majority of models explored supports the hypothesis that is a meaningful predictor of (near-)repeat victimisation (as well as the more general conclusion that space-time clustering is indeed influenced by the street network). The fact that the relationship does not appear to be universal, however, implies that this should be treated with caution and that further investigation may be required.

4.4.2 Implications for crime prevention and modelling

Given that crime is, in general, difficult to predict, the ubiquity of space-time clustering, coupled with the fact that its effects are concentrated in space and time, means that recent offending is one of the most powerful indicators of future crime available to the police. It is for this reason that (near-)repeat victimisations are among the crimes which are most feasibly preventable, via interventions which seek to interrupt or discourage the commission of the secondary offence.

Such interventions typically involve some targeted effort (*e.g.* patrolling or target hardening) in the vicinity of an initial incident; in theory, the spatio-temporal extent over which this is applied ought to be guided by what is known empirically about the places which will experience highest risk. On the basis of existing research,

however, the only terms in which this can be expressed is as an absolute Euclidean distance from an initially-victimised property: the assumption is that risk spreads in all directions.

The value of commonality in this context is that it allows the potential for criminality on the segments surrounding a victimised property to be assessed on the basis of factors other than just distance. The results presented here demonstrate that some streets are at significantly higher risk than others, and that these can be identified on the basis of network properties. This raises the possibility that interventions can be targeted preferentially towards areas of greatest threat, which ultimately represents greater efficiency when resources are limited.

One particularly relevant possibility is the use of mathematical models to guide the assignment of policing resources (this is the main topic of Chapter 5). Many of these employ some form of diffusion to model the spread of crime risk; however, the results presented here highlight one shortcoming of these. Whether situated in Euclidean space or on a network, such diffusion is taken to act uniformly in all directions, either radially or through the street network. The study of commonality shows that this not the case, and that the spread of risk in the real world is biased in certain directions.

The problem is particularly acute for networks, since the only natural way to formulate diffusion on a network is as a process acting between neighbouring links or vertices. This means that the spread is restricted spatially, and the effect of a disturbance is unlikely to be manifested substantially more than a few links away. Commonality, which provides a dyadic relationship between all links, could be used to act as some form of ‘conductance’ in such a model. In doing so, the street network would effectively be treated as fully-connected: all links would be connected to all others, though some connections would be very weak. In this way, longer-range transmission would occur naturally (and occur in such a way as to reflect commonal-

ity). This would, however, come at a cost: the commonality matrix does not have a simple mathematical characterisation, and so formal analysis would be intractable.

4.4.3 Wider applicability of commonality

Although motivated by a particular issue in the study of crime, the network metric commonality proposed in this chapter represents a significant outcome, independently of this particular application. Commonality extends the principle of the widely-used metric betweenness to a pair-wise relation by which the association between two network features, in terms of travel through the network, can be quantified. Such a concept is applicable in many areas of network science.

Betweenness is typically thought of as encoding two related ideas: the flow of traffic, and ‘brokerage’ in a network. For the first of these, applications in transport analysis can be envisaged in very similar terms to those explored in this chapter: questions of the form “if a vehicle travels through X, how likely is it to also travel through Y” are likely to arise frequently around issues of traffic management, for example. Extending this, a use could be foreseen in establishing the optimal placement of devices in order that they are encountered by the maximum number of unique vehicles (*e.g.* numberplate recognition systems, or even roadside advertising).

In terms of brokerage - typically a notion in social network analysis - possible applications are also apparent. For example, commonality could be regarded as a way of establishing roles within a network: if the commonality of two entities is high, it is likely that they have similar access to the remainder of the network and perhaps play similar roles in the transfer of information. In other fields in which social network concepts are used, such as in epidemiological modelling, it may provide a means of making inferences about the path of transmission of some infection through a network.

In summary, the ideas which underlie commonality are examples of more general

concepts which are applicable across a number of fields. Such possibilities may be investigated in future work.

Chapter 5

A model for burglary on street networks

5.1 Introduction

Having established various empirical results related to the distribution of burglary in time and space, the focus is now turned to the proposal of a mathematical model. As will be argued, modelling represents a natural means by which observational research can be used as the basis for tools with immediate relevance to both practice and policy. In addition, such models (and those of social systems in general) represent interesting cases from a mathematical perspective, since they can give rise to dynamics which are different from those observed for physical systems.

The empirical results presented in Chapters 3 and 4 have concerned the relationship between the street network and patterns of burglary, and such effects will remain the primary concern in the modelling work. The results of those chapters, together with those published elsewhere (*e.g.* Johnson & Bowers, 2010), provide support for a number of theoretical arguments which invoke the street network as a significant factor in shaping crime patterns. This implies that the absence of such effects from models represents a significant shortcoming; nevertheless, such considerations are absent from the majority of published work, and those examples which do exist are primarily agent-based in nature (*e.g.* Groff, 2007b; Malleson *et al.*, 2012). This work seeks to incorporate the effect of the street network in a new mathematical model.

The modelling process presents several mathematical challenges. The first of these is conceptual, and concerns the relationship between micro-level mechanisms and macro-level observations: relevant criminological theories relate to individual behaviour, but the phenomena to be modelled are collective and therefore require the aggregation of such behaviours. The recognition that not all such activities must be directly modelled, as well as the observation that network metrics can be used as heuristics for this, allows this to be achieved. The network aspect itself, however, also represents technical challenges. Situating models on complex topologies means that classical techniques - those which apply to continuous domains, for example - cannot be used, and solutions must be found by alternative means. This having been achieved, the challenge of interpretation also concerns the network aspect: how network structure influences model behaviour, and the implication of this for policy.

The chapter begins with a discussion of the relevance of spatio-temporal modelling to crime research, and a description of the various approaches which are typically employed. Existing work is then reviewed, including models which are not explicitly mathematical but which suggest approaches by which offender behaviour can be encoded formally. A new model is then described, and the associated dynamics characterised in terms of the underlying network structure. The features of the model are then explored both numerically and analytically, several of which correspond to issues of practical concern.

5.1.1 Modelling rationale

The question of how empirical research can be utilised in order to effect meaningful real-world outcomes is common to any field with relevance to policy, and is particularly pronounced when the issues concerned are social in nature. Research related to the distribution of crime is a prominent example of this: various phenomena have been observed consistently, and provide the basis for general theories of crime, but the optimal translation of these into practical recommendations remains an open problem. Such recommendations, however, represent a natural ultimate objective

for research in the area; indeed, a central tenet of ‘crime science’ (Laycock, 2005) is the necessity to retain this perspective throughout.

Policies motivated in this way can, of course, focus on different aspects of crime and concern various scales. For example, while concepts such as ‘hot-spot policing’ (Chainey & Ratcliffe, 2005) are concerned with the efficient deployment of police resources at relatively fine spatio-temporal scales, others provide much more general guidance related to planning and design in urban areas, such as the framework of ‘crime prevention through environmental design’ (Jeffery, 1971). Common to all cases, however, is that they rely on some hypothesised reasoning for why crime should occur in particular circumstances, and for the anticipated effect of a change in these circumstances.

When considered in this way, it is natural that crime is a field for which modelling approaches are likely to be profitably employed. The distillation of the theoretical mechanisms by which crime occurs to formal expressions not only adds clarity, but also facilitates quantitative analysis via the application of sophisticated mathematical tools. In the first instance, this provides a means for the testing of criminological hypotheses: to examine, in a robust and quantitative way, whether a given mechanism is necessary and/or sufficient to generate an observed pattern. Such results are of particular value in the criminal field, in which the increasing demand for ‘evidence based policing’ (Sherman & Eck, 2002) requires that principles should be quantitatively justifiable, rather than simply intuitive. Beyond this, analytical techniques can also be used to gain additional insight, by identifying latent interdependencies, for example, or by quantifying the relative effects of various factors on the distribution of crime.

The most significant appeal of models, however, lies in their ability to generate predictions. An example of this is the forecasting of future offending, which has attracted particular attention recently, and forms the basis for various approaches

generically referred to as ‘predictive policing’ (see Perry *et al.*, 2013, for an introduction). Bowers *et al.* (2004), for example, demonstrated the potential of a statistical method for this purpose, showing that its performance in prospectively identifying the locations of crimes was superior to common traditional approaches. Since then, variations and refinements of these methods have been developed commercially and procured by several police forces in various countries (Perry *et al.*, 2013). In these cases, model predictions generated by software are used to direct policing activity, particularly in the form of patrolling, aimed at both the prevention and detection of crime. With this outcome in mind, Groff & La Vigne (2002) emphasise the need for predictive methods to have clear theoretical grounding, since interventions can then be designed with specific consequences in mind.

In addition to simple forecasting, however, models can also be used to estimate the effect of a change in conditions or a hypothetical intervention. In the language of modelling, such variations correspond to a structural or exogenous change, and could represent either behavioural or environmental alterations. Assuming the validity of the model, its response to an intervention is a quantitative estimate, consistent with the hypothesised mechanism, of its effect; such an estimate can be regarded as the outcome of a rigorous form of thought experiment. Such *in silico* modelling represents a cost-effective aid to policy in many domains, since it is a risk-free means by which a very high number of hypothetical scenarios can be explored at low marginal cost. Again, crime is a field for which the value of such an approach is particularly pronounced, since the performance of real-world experiments is often problematic. Financial and ethical concerns represent significant constraints on such real-world studies, and indeed more fundamental concerns have been expressed regarding their utility and validity when applied to crime (Pawson & Tilley, 1997; Eck, 2002; Tilley, 2009; Knutsson, 2009).

Despite this motivation, it is only recently that the modelling of crime patterns has attracted interest within the academic community, for a number of reasons. The

first concerns data, and the fact that the availability of criminal incident records, in particular those with spatio-temporal information of sufficient accuracy, has historically been severely limited. Even if this had been available, however, the absence of software suitable for large-scale quantitative analysis of spatio-temporal data, together with contextual geographical data, also represented a significant barrier. From a more conceptual perspective, however, enthusiasm has also been limited by the prevailing view that the intricacies of social systems render them intractable by modelling methods which were originally developed for application to physical systems. The methodological innovations which constitute complexity science, however, address this aspect specifically, and their development has brought an attendant increase in modelling effort related to crime (*e.g.* Short *et al.*, 2008; Birks *et al.*, 2012). The increased interest within the mathematical community in particular is evidenced by the publication of a special journal issue on the topic (Berestycki *et al.*, 2010).

In terms of particular crimes, burglary is an especially appealing candidate for modelling, for practical and theoretical reasons. Firstly, it is a crime for which empirical phenomena are relatively well-known and well-defined: reporting levels are high (Budd, 1999), and concepts such as near-repeat victimisation have been demonstrated clearly and consistently (Johnson *et al.*, 2007). In addition, these can be rationalised with more general theories, such as routine activity theory and the notion of foraging, which provide useful grounding for model building. Finally, with the goal of such research in mind, burglary is a crime for which model outputs can be translated naturally into practical outcomes: by guiding interventions at the level of individual properties, for example, or by influencing urban planning.

5.1.2 Modelling approach

Before reviewing the crime modelling literature, it is appropriate at this stage to make a distinction between the several modelling approaches applied in this field, aspects of each of which will be employed in the model subsequently proposed.

Spatio-temporal models of crime can be assigned to three broad groups, each with distinctive merits: purely statistical approaches; those which use agent-based simulation; and more traditional mathematical models based on dynamical systems theory.

The term ‘purely statistical’ is used here to refer to those models which are essentially agnostic to generative mechanisms, drawing only upon numerical relationships. These models are constructed by examining the extent to which crime levels change with variation in certain explanatory factors, while accounting for random effects, and the relationships inferred can then be extrapolated for hypothetical values. Although this is certainly a rigorous process for data-driven estimation, the absence of a direct correspondence between the structure of the model and that of the system in question means that they offer limited theoretical insight; they discern the existence of relationships, but not *how* or *why* effects are manifested. This also implies that their applicability may be compromised in the event of a structural change, or change of circumstances, which may alter the relationships in a way which is unknown *a priori*.

The appeal of agent-based models (ABMs), on the other hand, lies precisely in their direct correspondence to hypothetical mechanisms; this is particularly appealing in crime, where the majority of theories are specified at the individual level. Since behaviours are encoded explicitly, these models relate most closely to emergent phenomena: those which arise from the accumulation of individual-level behaviours (Gilbert, 2008). In this sense, ABMs represent a ‘brute force’ approach, and are capable of incorporating arbitrary detail and granularity. Although this is advantageous to an extent, modelling is, in principle, concerned with simplification rather than total replication. The sense in which ABMs achieve this is by establishing *minimal* behavioural conditions for the occurrence of specified phenomena; beyond this, though, the insight that can be gained is limited to the statistical observation of simulations (which, at a sufficiently large scale for practical implementation, may be computationally prohibitive). In light of this, the most appropriate use of ABM

is therefore as a means of ‘proof of concept’ in the exploratory stages of model development, by which sufficient behavioural conditions can be established. In addition, simulation can reveal outcomes which were not envisaged in the construction of a model.

The remaining category, of mathematical models in the more classical sense, necessarily encompasses a wide range of possibilities. Typically, though, such models will consider the environment and the actors within it as a dynamical system, and derive differential equations on the basis of hypothesised mechanisms. Although this process is reductive by nature, the move to generality does allow for the analysis of such models using established mathematical techniques. The questions which can be addressed using these include those related to stability and to the sensitivity of the systems to changes in parameters or conditions; more generally, they can be used to identify the fundamental drivers of macro-level behaviour. Crucially, observations can be made *a priori*; that is, before formal solutions are found. As with physical systems, insights can be translated back into terms of the system described by the model and, ultimately, their theoretical implications explored.

Of course, the distinction between these approaches is not a rigid one, and several examples incorporate elements of more than one. In certain (sufficiently simple) cases, for example, observable characteristics of an agent-based model can be described by differential equations, and this duality means that the approaches are, to an extent, interchangeable. In addition, those examples which model the symptoms of criminal activity by analogy with some physical system (*e.g.* diffusion of risk) can be regarded as both statistical and mathematical with good justification.

5.2 Review of previous modelling approaches

Although the focus here is on mathematical modelling, with application specifically to burglary, the body of relevant literature is somewhat more broadly defined. Since many of the concepts involved are not unique to this crime type, or approach, the

scope here includes all modelling of the spatio-temporal distribution of crime. The model of burglary proposed by Short *et al.* (2008), which has attracted particular attention and which corresponds closely to the model introduced here, will be reviewed in particular depth.

Statistical models

One of the earliest attempts to consider the spatio-temporal distribution of crime in a formal modelling context was published by Gunderson & Brown (2000), who presented a framework for the agent-based modelling of both physical and cyber crime. A notable feature of the approach proposed is that it involves the use of machine learning methods as a means of establishing the parameters for a subsequent agent-based simulation. It was suggested that, given a set of crimes, clustering methods could be used to infer the number of criminal actors responsible, and that inferences could then be made about the criminal preferences of each by considering the features of their targets. These could then form the input for a multi-agent simulation which could be used for predictive purposes; no practical implementation, however, was presented.

The agent-based component was not present in subsequent work (Liu & Brown, 2003), which focussed only on the use of machine learning for prediction. In the model presented, a set of environmental features is extracted for each of a known collection of incidents, and machine learning methods then used to find a subset of ‘key features’ which best explain the original point pattern; these are then used to build a predictive statistical model. When applied to data for breaking and entering in Richmond, Virginia (US), the method out-performed kernel density methods based on the point patterns alone, though the dataset considered was small. The model in question is entirely data-driven, and its treatment of spatial and temporal distributions as independent is not supported by empirical findings elsewhere (*e.g.* Johnson *et al.*, 2007).

Some of these concerns were addressed in subsequent similar work. Xue & Brown

(2006) incorporated criminal-decision making by adapting a discrete spatial choice framework, the inputs to which were again determined using a clustering method. This model also out-performed kernel density methods when applied to Richmond data. A later approach (Wang & Brown, 2012) then used generalised additive models as a means of combining spatial and temporal information, including, for example, a ‘time since last incident’ variable. Again, this performed better than traditional methods when applied to data for breaking and entering in Charlotte, Virginia (US).

A number of other statistical models use the concept of event-dependency as their basis. Bowers *et al.* (2004) proposed a method, ‘ProMap’, in which an urban area is divided into grid cells, and the prospective risk for each of these is estimated using recent crimes which have occurred nearby. Each incident contributes a value which is inversely proportional to both the time since it occurred and the number of cells’ separation, reflecting the decay in both space and time of the elevation in risk. This was shown to out-perform standard methods for long-term hot-spot identification when identifying burglary incidents in Merseyside, UK. A subsequent paper (Johnson *et al.*, 2009a) refined the approach by incorporating a multiplicative term to account for heterogeneity of opportunity, leading to a modest improvement in predictive accuracy.

The same principle as ProMap is used in the self-exciting point process model of Mohler (2011), who drew an analogy with patterns of earthquake occurrence. The primary innovation in this case was the use of a stochastic de-clustering algorithm to estimate the relative contributions of background offending and self-excitation in a set of incidents; for each offence, a set of probabilities are inferred for whether it was triggered by each of the other offences. In this way, both effects can be estimated from data, and the method out-performed ProMap on burglary data for Los Angeles.

One further statistical approach was proposed by Rey *et al.* (2012), who analysed burglary in Mesa, Arizona using conditional spatial Markov chains. Again using a

grid system, the process described examines whether the transition of a given cell between states ('burglary' or 'no burglary') from one time unit to the next is influenced by the states of its neighbouring cells. This was found to be the case, as would be expected, and, although no attempt was made to use it predictively, the general approach is a viable one.

Agent-based models

An early example of an agent-based approach is the model of street robbery proposed by Liu *et al.* (2005). Although presented as a cellular automaton, the model is essentially an agent-based one, and the cellular aspect simply defines the environment in which agents move. The model takes into account routine activities, and is also notable for its formulation of 'tension' as a property of the environment which is elevated by the occurrence of crimes and which diffuses across space. Simulated results appear to broadly reproduce known features, such as repeat offending and victimisation.

A subsequent book by Liu & Eck (2008), which explored the use of simulation in criminology much more widely, features a number of studies on this theme. Hayslett-McCall *et al.* (2008) proposed a similar model to Liu *et al.* (2005), though adapted for the crime of burglary and using real offender data for Dallas, Texas. The motivation in this case was investigation of the journey to crime, though the simulation results were not investigated in depth. A rather more basic question is addressed by Brantingham & Tita (2008), who explored the types of victimisation patterns produced by two simple behavioural principles: movement modelled as a Lévy flight process, and spatial anchoring. Only single agents are considered, and the environment is an abstract, homogeneous 2-dimensional space, but simulation results indicated that even simple principles such as these are sufficient to generate intricate spatial patterns.

At the other extreme of complexity, the modelling framework proposed by the Mastermind model (Brantingham *et al.*, 2008) aims to produce an extremely detailed

and theoretically comprehensive modelling environment. A central principle of the work is its modular structure, by which separate elements of the model (*e.g.* route choice, awareness spaces) are handled by dedicated routines, using an abstract state machine formalism. This allows individual elements of the model to be refined, up to an arbitrary level of detail, and therefore represents a highly versatile approach. Navigation, which takes place on street network graphs, is determined by shortest paths (although the definition of distance can incorporate various subjective factors), and the variation in these preferences appears to have a material effect on the simulated behaviour.

A similarly detailed approach is taken by Malleson *et al.* (2010), who describe a highly-detailed model for burglary, designed to be run in a realistic spatial environment. The model uses the PECS (Physical conditions, Emotional states, Cognitive capabilities and Social status) framework for agent simulation, and focuses in particular detail on the mental state of criminal agents, incorporating their needs for money, drugs and sleep. Routine activities are encoded explicitly, and awareness spaces are stored in memory as agents travel between activity nodes (which they do via shortest paths). Burglary events themselves are a two-stage process, involving area-level identification followed by local exploration, and the paper examines the effect of crime prevention strategies on patterns of offending. In a later paper (Malleson *et al.*, 2012), real geographical data for the city of Leeds, UK, is used in conjunction with data on known offenders to examine its predictive power: the general patterns observed are well matched, although significant discrepancy in individual behaviours still remains.

While undoubtedly offering the closest correspondence with reality, elaborate models such as these are not well-suited to the testing of particular criminological hypotheses, since particular effects may be difficult to identify amid such complexity and variability. For this reason, other approaches are simple by design, and adopt a pseudo-experimental approach by making individual well-specified changes to the

model while keeping all other factors constant. An example of this is the model of Johnson (2008), which is used to examine the sufficiency of the ‘flag’ and ‘boost’ explanations for repeat victimisation in burglary. Although not strictly agent-based (in the sense that offender agents are not simulated explicitly), the approach simulates offences using real geographic and demographic data, finding that both mechanisms appear to play a role in generating patterns of victimisation.

A similar principle is used by Birks *et al.* (2012), who present a unified framework for examining the role in shaping burglary victimisation patterns of three key concepts: routine activity theory, the rational choice perspective and crime pattern theory. Each of these is codified as a set of rules for agent behaviour, and a series of simulations are carried out in which each factor is ‘switched’ on or off until all combinations have been exhausted. The goodness of each output is assessed by measuring three features of the simulated victimisation patterns - spatial clustering, repeat victimisation and the journey-to-crime - the existence of each of which can be regarded as a benchmark for generative models. Though their contributions are manifested differently, the addition of each theoretical principle appears to represent a positive refinement, and in combination they appear to be ‘generatively sufficient’ to reproduce the three stylised patterns discussed.

The work of Groff (2007a,b) is similar in approach, but is of particular relevance in the present context due to its emphasis on street network effects. The main focus of the work is on testing the impact of routine activity behaviour on the incidence of street robbery, but it is argued that such activities are influenced so strongly by urban form that they cannot be properly understood without constraining movements to a realistic street network. Indeed this proves to be the case, as results differ significantly when street data, imported via GIS, are used instead of a grid. Although no attempt is made to relate the results to network structure, this does highlight the sense in which the effects of certain mechanisms for crime depend crucially upon environmental factors.

A contrary view is offered by Elffers & Van Baal (2008), who argue that the inclusion of a realistic spatial backcloth in models of crime is counter-productive. They suggest that doing so is of primarily cosmetic value, and does not add to theoretical understanding. In fact, they suggest that the irregularities introduced by idiosyncrasies in geographical data can mask effects, so that theoretical insights may be more difficult to identify. Of course, this argument is no longer valid when the effects in question are explicitly determined by the spatial backcloth, *i.e.* its influence is precisely what the research seeks to examine. Nevertheless, a related point concerning external validity still stands: backcloth should be incorporated in a generic manner, and not be applicable only in a particular geographical setting.

The street network is also a central component of the CriMM model proposed by Iwanski *et al.* (2011). This model builds on empirical research concerning directionality in the journey to crime (Frank *et al.*, 2012) by inferring from an offender's offence location (relative to his or her home) that he or she was in the process of travelling to a certain crime attractor (see Brantingham & Brantingham, 1995). It is then shown that, if that journey is taken to be a shortest path through the network, the offences show a clear distance decay pattern away from that path. A later paper (Iwanski *et al.*, 2012) explores the opposite problem, of whether extrapolating the routes taken for journeys-to-crime beyond the offence location tends to lead to locations that would be classed as crime attractors. The way in which paths are extended is determined by a heuristic which combines directionality and a measure of network activity similar in spirit to betweenness, and indeed these do tend to lead to likely crime attractors. It is worth noting, however, that the offender journeys modelled in both these cases will tend towards higher-betweenness streets by construction, and so the results are consistent with the rather simpler hypotheses that: a) crimes tend to occur near high-betweenness streets; and b) crime attractors tend to be located on high-betweenness streets.

5.2.0.1 Mathematical models

The model proposed by Berestycki & Nadal (2010) is, in contrast to the above, explicitly mathematical in its construction. The model is situated on an abstract continuous space and describes the evolution in time and space of generic criminal activity, taking into account several aspects of criminological theory. The fundamental quantity in the model is a ‘willingness to act’ field, which encapsulates all features which contribute to the propensity of a location to experience crime at a given time, such as the availability of offenders, suitability of targets and strength of deterrence.

A crucial innovation, however, is the use of a threshold function for the *level* of criminal activity as a function of the willingness to act. The use of a threshold means that the model takes account of the fact that, even though crime may be absent in an area, this encompasses a wide range of possibilities for the criminal propensity: it may be extremely low, or only marginally lower than the threshold. Deterrence is incorporated via an adaptive cost function, and various social influences included as diffusive terms. Mathematical analysis of a simple form of the model reveals several interesting phenomena, including a characterisation of two equilibrium states in space. ‘Warm spots’ are observed where propensity is high, but is constrained by large deterrent activity, and ‘tepid milieu’ are characterised by a low, but non-zero, level of criminality pervasive in space, both of which describe situations observed in the real world.

Berestycki *et al.* (2013) performed further analysis of this model, considering in particular manipulation of the population’s natural tendency towards crime. It was demonstrated that certain configurations can give rise to travelling wave solutions, which have real-world interpretation as criminal invasions across space. This is a finding with practical implication, since the problem of control (*i.e.* policing) is considerably different in this case. This issue was addressed specifically, and a minimum level of resource required to prevent the onset of criminal invasion was established.

5.2.1 The Short model

The model of burglary introduced by Short *et al.* (2008) is one which has attracted particular attention, and has inspired several follow-up analyses concerned with both theoretical refinement (*e.g.* Pitcher, 2010) and formal mathematical analysis (*e.g.* Kolokolnikov *et al.*, 2014). In addition, it is distinguished by its close correspondence with theories specified at the level of the individual, since the mathematical expressions are derived from behavioural principles. As such, it is a prototypical example of how a model formulated in agent-based terms can be translated to a continuous form, and exemplifies the duality of the two approaches. In light of the compelling case made within previous research for the value of the individual perspective, grounding models in this way is prudent from a theoretical point of view (and ought to render them more persuasive). Because of its relevance to the work presented here, the model is described in detail below.

5.2.1.1 Discrete form

The model is first formulated in discrete terms and describes the behaviour of a simple urban system. The system is composed of two types of entity: houses, at which burglaries can occur; and burglars, who seek to commit them. The houses are arranged on a finite two-dimensional regular lattice, with exactly one house at each lattice site. The houses are indexed by s and each is associated with a strictly positive value, A_s^t , which describes the *attractiveness* of property s at time t , as perceived by a prospective burglar. Motivated by the ‘flag’ and ‘boost’ theories of repeat victimisation, the attractiveness A_s^t is composed of a static component, A_s^0 , and dynamic component, B_s^t (again both strictly positive):

$$A_s^t = A_s^0 + B_s^t. \quad (5.1)$$

A_s^0 is intended to encapsulate all time-stable factors which affect a property’s suitability as a target (*e.g.* affluence or ease of access), whereas B_s^t represents any supplementary risk to which the property is subject (such as that associated with a

recent victimisation).

The lattice is populated by burglars, each of which is located at some point on the lattice; N_s^t is taken to be the number of burglars at location s at time t . The system evolves according to a discrete time process, in which the actions of burglar agents and the state of the environment are updated at regular intervals of length δt .

At each time-step, burglars decide whether to commit a burglary at their current location; this is modelled as a Poisson process with rate given by the attractiveness, A_s^t , of the location. With δt sufficiently small, the probability, p_s^t , of an event at s in a given time interval $[t, t + \delta t)$ can be approximated as

$$p_s^t = 1 - e^{-A_s^t \delta t}. \quad (5.2)$$

If the burglary does take place, the burglar is immediately removed from the system; otherwise, the burglar moves to one of the 4 immediately neighbouring houses according to a random walk. This movement is biased towards more attractive locations, so that the probability of movement from a location s to a neighbour w is given by $\frac{A_w^t}{T_s^t}$, where $T_s^t = \sum_{s' \sim s} A_{s'}^t$ gives the total attractiveness of the neighbours of s (the relation $s' \sim s$ indicates that s' is an immediate neighbour of s).

In addition to the actions of existing agents, each time-step also includes the addition of new burglars. These are created at each site at a constant rate of Γ per unit time, and the population of burglars is therefore constantly varying as some are removed and others are generated.

The dynamic attractiveness B_s^t is also updated at each time-step, and it is at this stage that the ‘boost’ effect of previous victimisations is taken into account, in two ways. First, each burglary causes the dynamic attractiveness of the burgled house to be instantaneously increased by θ ; however, given the finite lifetime of such an effect, this decays back to 0 over time. Secondly, to account for the near-repeat

phenomenon, dynamic attractiveness also diffuses across space, from houses to their neighbours. The evolution of B_s between times t and $t + \delta t$ is therefore given by

$$B_s^{t+\delta t} = \left((1 - \eta)B_s^t + \frac{\eta}{z} \sum_{s' \sim s} B_{s'}^t \right) (1 - \omega_1 \delta t) + \theta X_s^t, \quad (5.3)$$

where X_s^t is a random variable representing the number of burglaries which occurred at s during $[t, t + \delta t)$, η is a diffusion parameter, ω_1 is the inverse mean lifetime of the dynamic effect, and z is the number of neighbours of s (uniformly 4 in the case of the lattice).

The system described is a stochastic one, with randomness present in agents' movements and decisions to offend, but a corresponding deterministic formulation can be derived by replacing X_s^t with its expected value, $N_s^t p_s^t$. This discards the notion of burglars as distinct entities and constitutes a 'mean-field' model for their averaged behaviour. The dynamic attractiveness can therefore be expressed in terms of N thus

$$B_s^{t+\delta t} = \left((1 - \eta)B_s^t + \frac{\eta}{z} \sum_{s' \sim s} B_{s'}^t \right) (1 - \omega_1 \delta t) + \theta N_s^t p_s^t. \quad (5.4)$$

A difference equation for N itself can also be written by observing that the expected number of agents at a location s that do not offend (and therefore move) is given by $N_s^t(1 - p_s^t)$. Using this in conjunction with (5.4) and the 'birth rate' Γ gives the relation

$$N_s^{t+\delta t} = A_s^t \sum_{s' \sim s} \frac{N_{s'}^t(1 - p_{s'}^t)}{T_{s'}^t} + \Gamma \delta t. \quad (5.5)$$

At this point, it is possible to note a steady equilibrium solution to the system for a simple case: that where the static component of the attractiveness is given a uniform value \tilde{A}^0 across all sites. When this is the case, a spatially-homogeneous equilibrium for uniform values \bar{B} and \bar{N} can be found. The solution is found by considering that, for this to be the case, the decay in attractiveness in each time-step must be equal to the increase due to further burglaries:

$$\omega_1 \bar{B} \delta t = \theta \bar{N} \bar{p}, \quad (5.6)$$

and also that the number of agents which leave the system due to burglary must be balanced by those which are generated:

$$\overline{N}\overline{p} = \Gamma\delta t. \quad (5.7)$$

These can then be used to solve for the equilibrium values

$$\overline{B} = \frac{\theta\Gamma}{\omega_1} \quad \overline{N} = \frac{\Gamma\delta t}{1 - e^{-(\tilde{A}^0 + \overline{B})\delta t}}. \quad (5.8)$$

5.2.1.2 Numerical results

Short *et al.* began their analysis by presenting the results of numerical simulations for an implementation of the original agent-based form. The simulations were carried out on a 128 x 128 grid and initialised with $B_s^0 = \overline{B}$ for all s and \overline{N} agents randomly distributed throughout the lattice. As the parameter configurations were varied, three apparent régimes were observed for the attractiveness field:

- a) Spatial homogeneity, in which fluctuations from the mean value caused by burglaries dissipate quickly.
- b) Stationary hot-spots, where the system tends towards a steady state in which well-defined circular areas of high attractiveness are surrounded by areas of low attractiveness. The size and configuration of these is parameter-dependent.
- c) Dynamic hot-spots, in which identifiable hot-spots form and exhibit spatial drift. These either persist in this dynamic state or disappear after varying lengths of time.

The authors suggest that, in fact, simulations which fall into régime (c) are actually manifestations of (a) or (b), where the drift phenomenon is an artefact of finite size effects. Configurations which fall into régime (c) are those which involve relatively low burglar numbers, implying that perhaps the low levels of stochasticity, relative to the size of the system, are the reason for the phenomenon. More specifically, the authors suggest that the number of agents acting in the system is insufficient

for mean-field averaging to cancel the effect of stochastic fluctuations upon the aggregate patterns. Hot-spot formation characteristic of regime (b) is reproduced, for illustration, in Figure 5.1.

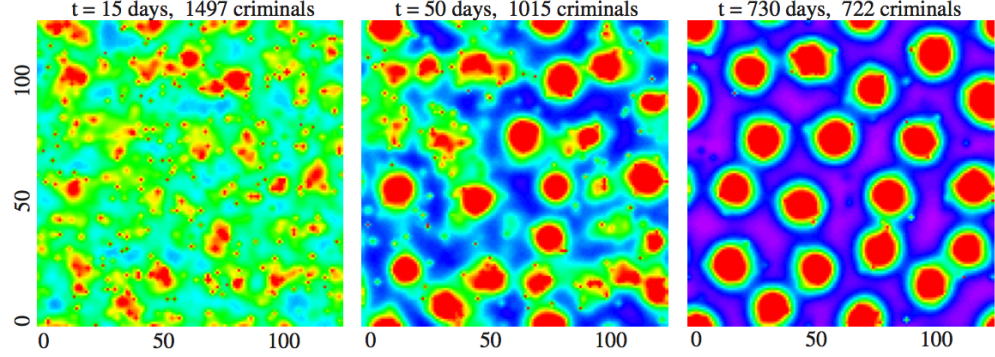


Figure 5.1: Formation of static hot-spots in a discrete simulation of the model of Short *et al.*. The colours represent the dynamic attractiveness field, B_s , with colours ranging from blue (low) to red (high).

5.2.1.3 Continuum limit

The final component of the analysis presented by Short *et al.* concerns the derivation of a continuum limit for the model, after which the resulting PDEs are analysed formally. As will become clear, the fact that this chapter is concerned primarily with modelling crime in discrete space means that such analysis is not expressly relevant in the present setting, and so this aspect of the analysis will not be described in full. The resulting expressions, however, can be briefly summarised. Having replaced the terms A_s , B_s and N_s with equivalent continuous fields A , B and N respectively, the system is encapsulated in the two equations

$$\frac{\partial B}{\partial t} = \frac{\eta D}{z} \nabla^2 B - \omega_1 B + \theta N (A^0 + B) \quad (5.9)$$

$$\frac{\partial N}{\partial t} = \frac{D}{z} \nabla^2 N - \frac{2D}{z} \nabla \cdot \left(\frac{N}{A^0 + B} \nabla (A^0 + B) \right) - N (A^0 + B) + \Gamma, \quad (5.10)$$

where D defines the speed of movement of burglars. Standard boundary conditions are given by

$$\left. \begin{array}{l} B = B_0(x, y) \\ N = N_0(x, y) \end{array} \right\} \quad t = 0, \forall (x, y) \in \Omega, \quad (5.11)$$

for a domain Ω . These boundary conditions, however, may be varied in order to reflect alternative scenarios: periodic conditions would approximate a large spatial area, whereas no-flux conditions might imply an urban area with limited spatial extent. Solving for a spatially-homogeneous equilibrium is straightforward, and gives:

$$\bar{B} = \frac{\theta\Gamma}{\omega_1} \qquad \bar{N} = \frac{\Gamma\omega_1}{\theta\Gamma + A^0\omega_1}. \qquad (5.12)$$

With the model in this form, linear stability analysis is then carried out in order to characterise the conditions under which the spatially-homogeneous solution is unstable, and for which the model will thus give rise to hot-spots. In fact, the condition which is ultimately found is one which relates the ‘area of influence’ of any individual boost to the average number of events per area at steady state. It is therefore argued that distinct hot-spots of risk can only occur if they are sufficiently well-separated in space so as not to interact.

5.2.1.4 Further developments

The model of Short *et al.* has inspired a number of subsequent papers, comprising both sophisticated mathematical analysis and variations to the model itself. In the former case, the contributions are primarily technical (in that the model is not developed from a criminological perspective); however, as noted above, these are of limited relevance in the context of a network-based model, so will be discussed only briefly. A prominent example is the analysis of Kolokolnikov *et al.* (2014), in which the question of stability is translated to an eigenvalue problem, and in which conditions are established for the stability of patterns with a fixed number of hot-spots. Cantrell *et al.* (2012) demonstrated that the system supports global bifurcation, and that the resulting spatially-heterogeneous patterns can be stable near the point of bifurcation. Finally, Lloyd & O’Farrell (2013) extended the analysis of Short *et al.*, showing that localised hot-spots can bifurcate from homogeneous solutions, and tracing their development via path-following.

Other developments to the original model have, on the other hand, incorporated substantially more theoretical innovation. In particular, these have included the addition to the model of the effect of policing. The first of these was proposed by (Jones *et al.*, 2010), and incorporated both the influence of policing on criminal activity and the movement of officers. For the former, two mechanisms were included: the presence of police officers diminished the attractiveness of individual houses, and prospective offenders also had a certain tendency to simply leave the system when they encountered an officer. Three possibilities for police movement were considered: it was modelled as either an entirely random walk, a random walk biased towards higher attractiveness, or a movement towards the periphery of hot-spots. Using both continuous and agent-based formulations, it was found that the strategy biased towards higher attractiveness was generally successful in suppressing crime, but that the strategy aimed at the periphery was also beneficial for hot-spots of larger area. Random patrolling was found to be ineffective.

Pitcher (2010) also incorporated police activity, alongside a number of other refinements. In that paper, it was argued that the fact that burglars could only be removed from the system was unrealistic, and had the counter-intuitive effect of increasing the burglar population when the level of crime was reduced (since they are generated at constant rate). It was suggested that, rather than searching eternally until an offence was committed, burglars may be removed from the system due to ‘fatigue’ if they have failed to find a target within a certain time. On the other hand, motivated by theories of repeat victimisation, the assumption that burglars leave the system after an initial offence was removed. This was achieved by simply including a decay term in the offender movement model, representing the ‘lifetime’ of a burglar.

The policing component of Pitcher’s model is realised via a *deterrence* field, which has the effect of causing a proportional decrease in attractiveness. Since attractiveness drives offender movement, this also has the effect of biasing offender movement

away from areas of high deterrence. Importantly, though, deterrence is not simply proportional to the number of officers; rather, it is an effect which spreads over time from officer locations, encoding a ‘diffusion of benefit’ associated with visible policing (see Clarke & Weisburd, 1994). Officer locations themselves are determined exogenously, but are typically updated periodically by moving officers (the number of which is fixed) to areas of highest attractiveness. Using simulation, it was shown that this deterrent effect was capable of both moving hot-spot locations, or deforming them to eventual ‘ring’ shapes.

One final piece of related work of note is another paper by Pitcher & Johnson (2011), which concerns an agent-based implementation of the modified model, though with the policing component removed. In addition, a ‘carrying capacity’ term was added, in order to prevent attractiveness reaching unrealistic levels. The purpose of the analysis was to examine whether this model, in which incidents occur explicitly, was capable of producing patterns of (near-)repeat victimisation similar to those observed empirically. After varying parameters, results suggested that the boost mechanism present in the model was capable of producing realistic effects; the flag account only, however, was not. Although this does not rule out other potential mechanisms, it does show that the capability of the model is not limited to the formation of spatial hot-spots.

5.2.2 Evaluation of the Short model

The results of Short *et al.* and Pitcher have demonstrated that the general framework described here is sufficient to generate certain well-known features of burglary offending patterns, such as hot-spot formation and repeat victimisation. In particular, it appears that diffusion-like processes are well-suited to the modelling of risk communication across space. In the original model and its derivatives, however, the spatial environment is modelled in a very simple way, as a two-dimensional regular lattice of houses (although this interpretation no longer holds in the continuum case, it remains the theoretical basis). Several assumptions concerning urban spaces are

implicit within this. First, it suggests that houses are laid out regularly in space, which is not the case in real urban areas (particularly in the UK), but is a simplification which does not pose particular conceptual difficulties. More problematic, however, are the implications for both offender movement and risk diffusion. In the model, burglars can move, and perceive attractiveness, uniformly in all directions, whereas these would both be constrained by the street network in a real setting. Similarly, the notion that the spread of risk is uniform in all directions is not consistent with the hypothesis that transmission is more likely between places which tend to lie in the same activity spaces (see Johnson & Bowers, 2007, for further discussion). If the travel distance between two places is high, there is no reason why a victimisation at one should influence the risk at the other, even if the Euclidean distance is small. Indeed, empirical work presented in this thesis and elsewhere (*e.g.* Johnson & Bowers, 2010) provides clear motivation for the incorporation of street network effects.

In the discrete case, adapting the Short *et al.* model to the case of an arbitrary network is straightforward: instead of the two-dimensional lattice, the ‘neighbourhood’ of each house (for the purposes of burglar movement and risk diffusion) is specified by the network structure, with all other aspects unchanged. By specifying a network which represents the arrangement of houses on streets, burglar movement can then only take place between houses for which pedestrian travel is possible. Although the continuum limit is no longer tractable (or, indeed, meaningful) in this case, simulations could be carried out to explore how the previous results concerning hot-spot formation and repeat victimisation are influenced by network variation. While this would, of course, generate interesting results, several more fundamental concerns suggest that a novel approach may be more appropriate.

Firstly, while all arguments for situating a burglary model on a network invoke, at some level, ideas from routine activity theory, they are absent from the Short model. The (memoryless) random walk mechanism by which agents move is, by

definition, incompatible with the existence of an activity space and a tendency to offend in certain areas. In addition, such a choice removes the possibility of examining other environmental effects; the influence of crime attractors, for example, is predicated on the idea of purposeful movement by offenders. Of course, to encode such concepts in a mathematical model such as this would represent a considerable challenge, if not an impossibility. Deriving an equation for the number of burglars at each location assumes that all are governed by the same behavioural rules, whereas this is not the case: individuals will decide their next movement differently depending on whether they are at home or at work, for example.

A further issue with the random walk hypothesis is the fact that it is biased towards areas of higher dynamic attractiveness. This draws the population of offenders towards existing ‘hot’ locations and serves to reinforce them via repeated offending. Modelling movement in this way, however, implies that dynamic effects apply to all agents, *i.e.* that they are all attracted to the vicinity of previously-victimised locations. In effect, agents possess ‘global’ information about the attractiveness of properties, which is theoretically problematic. There is strong evidence that boost effects are primarily a manifestation of same-offender victimisation (Bernasco, 2008; Johnson *et al.*, 2009b): an optimally-foraging offender (see Johnson & Bowers, 2004b) returns to a property because the successful execution of a burglary renders it a more appealing target. There is no reason why many of the factors which contribute to this (*e.g.* increased knowledge of how to access the property and of the goods available inside) should apply to any other offender. Even if a previously-victimised property was appealing for other reasons, it is extremely unlikely that the fact of its victimisation (or that of one of its neighbours) could be established by visual inspection. The boost effect therefore ought to apply to a single agent - the one who committed the triggering offence - and not influence movement universally. Again, this calls into question the mechanism by which hot-spots are generated in the model of Short *et al.*.

5.3 Proposed model description

The aim of this section is to propose a mathematical model for residential burglary, with two primary motivations: to incorporate street network effects by design, and to address the conceptual concerns associated with existing models. The reason for seeking a model based on differential equations is the wish to retain close correspondence with criminological theory while taking advantage of the parsimony and analytical possibilities afforded by such an approach. As argued in Section 5.1.2, modelling in this way represents an effective compromise between realism and analytical power, and is perhaps most easily deployed in practice.

Nevertheless, several aspects of the model are grounded in ideas from both statistical and agent-based approaches. Since the main phenomena of interest - the temporary elevation of risk associated with burglary incidents, for example - can only be meaningfully quantified in statistical terms, their representation necessarily has its basis in statistical modelling. Likewise, since the relevant theories of criminological behaviour describe the decisions of agents, one of the objectives of the model is to incorporate such effects in a well-motivated way. In particular, the inclusion of stylised notions of spatial behaviour (*e.g.* routine activities, activity spaces) and their effect on crime is of primary concern. While encapsulating the aggregated effect of idiosyncratic activity patterns using differential equations is clearly difficult, it is argued here that it is not, in fact, necessary to model such behaviour dynamically. Instead, such effects can be captured via the structure and spatial setting of the model itself; this represents the key step in capturing individual-level effects. In this way, the structure of spatial interactions occurring in an urban setting becomes a model in itself, which then determines the behaviour of the dynamic model.

5.3.1 Spatial setting

The setting for the model is a graph, $G = (V, E)$, which represents an urban street network. As with the empirical work presented in Chapters 3 and 4, the *primal* rep-

representation of the network is used, in which vertices are placed at street junctions, and a link exists between any two vertices for which the corresponding junctions are directly connected by a street (this representation and others are described in detail in Section 3.1.4).

At various points in the remainder of this chapter, model behaviour will be demonstrated using an example of a real-world network. The example used is a section of the street network of Liverpool, UK, the construction and representation of which is shown in Figure 5.2. This section includes a number of features which are convenient for the purpose of illustration, but the choice is otherwise arbitrary.



Figure 5.2: A section of the street network of Liverpool, UK: a) the original map, and b) the overlaid primal network representation, with vertices shown in red and links in black.

The notation used will remain the same as that used previously: the graph is taken to have n vertices, labelled v_1, \dots, v_n , and m links, labelled e_1, \dots, e_m . When they are used without indices, v and e will be taken to represent generic vertices or links. In addition, all links are encoded in the adjacency matrix $\mathbf{A} = (a_{ij})$, which is defined

so that

$$a_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are connected by a link} \\ 0 & \text{otherwise.} \end{cases} \quad (5.13)$$

Defining the graph in this way implies that the fundamental spatial unit in the model is the street segment. This differs from previous work (Johnson, 2008; Short *et al.*, 2008; Pitcher & Johnson, 2011), which has been based on a structure composed of individual houses. Such a structure could, of course, be used here, with adjacency defined on the basis of the street network, and Figure 5.3 illustrates one way of doing this. In this case, houses (represented by vertices) are first associated with the closest point on the street network to which they lie, and links are then added for any pair of houses between which an uninterrupted path exists via the street network. While such a construction would facilitate the modelling of more fine-grained phenomena, it is not used here for several reasons.

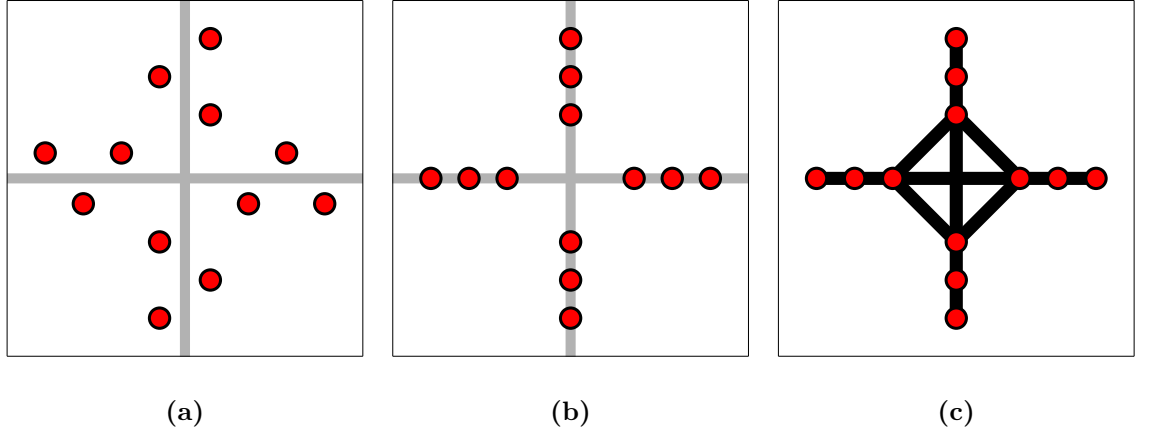


Figure 5.3: Construction of a network of houses: a) a hypothetical example of two intersecting roads (shown grey), with houses situated on either side; b) each house associated with the closest point on the street network; c) network links added between any pair of houses connected by an uninterrupted section of road.

Firstly, since the primary motivation for this work relates to effects at the street network level, the inclusion of additional granularity is analytically unnecessary: there is no reason to anticipate that ‘intra-segment’ patterns should be related to network structure. On the contrary, their inclusion has the potential to obfuscate larger-scale

patterns while offering little insight. Although significant within-segment patterns may exist, and could be modelled, it is at the segment level that model outcomes are likely to be of most practical use (in guiding police interventions, for example).

The second reason is technical, and follows a similar argument to that presented in Section 3.2.1.1. For any path-based network metric, all houses on a given segment take exactly the same value: this is because navigational choices occur only at junctions, so that trips must only use entire segments. For two adjacent properties on a given segment, any journey which passes one must also pass the other (since there is no opportunity to turn), and so they are indistinguishable on that basis¹. Again, therefore, the inclusion of individual houses adds little extra detail.

For these reasons, the model is constructed with the street segment as the basic spatial unit. Having defined the structure in this way, the ‘victimisation’ of a street segment therefore represents, in real terms, a burglary occurring at any of the houses which lie on it. A series of victimisations on a given segment may actually take place at different houses, but that is not relevant in the present scope.

5.3.2 Modelling approach

When building a model for which one of the fundamental underlying principles is the coincidence in space and time of offenders and targets, a natural approach (and that taken by Short *et al.*) is to incorporate a detailed dynamic model of offender movement. Indeed, the fact that the influence of the street network is expected to be exerted primarily via the shaping of such movements appears to further support the use of such an approach. However, scrutiny of the underlying theoretical mechanism suggests that this may not, in fact, be necessary. This can be seen by reviewing the two ways in which the movement of individuals is likely to influence the spatial distribution of crime.

¹Of course, the definition of betweenness could be modified in this case to count journeys between every pair of houses, rather than every pair of junctions. In this case, neighbouring houses would not have exactly the same betweenness value, but they would differ by at most 1.

The manner in which the movements of individuals determine their activity spaces, and therefore their awareness of opportunities for crime, is one of the central concerns of the model. Since some places feature in the activity spaces of more people than others, this implies that some areas will experience greater crime because the opportunities within them are encountered more often; an (uneven) opportunity surface exists across the network. However, the shape of this surface remains, for all practical purposes, constant at the time-scale over which the effects captured in the model act (time-of-day variation is beyond the present scope). At the very least, it is not influenced by recent offending - burglaries do not, in general, change the routine activities of the population at large - and so there is no reason for it to be coupled with 'real-time' crime activity.

The role of activity spaces in near-repeat victimisation is another mechanism by which the influence of the street network is exerted. As argued in Chapter 4, it is to be expected that, having committed one burglary, an offender is more likely to commit a follow-up offence elsewhere in his/her activity space. Since that space is shaped by movement patterns, there is reason to expect that the spatial distribution of risk elevation can be reconciled with network topology, and the results of Chapter 4 support this. Again, though, this is a static effect: the spatial relationships which determine how risk might be communicated exist independently of the burglary event itself. That is, the places to which risk is likely to diffuse, given an initial incident on a given street, are known *a priori* and do not change over time.

Given the observations in the previous two paragraphs, there is little reason why offender movement need be modelled dynamically; instead, those aspects which have relevance to dynamic patterns of crime can be represented in both the static properties of the network and the manner in which risk is communicated. The effect of individual offender movement is emergent, but it emerges independently of ongoing offending.

5.3.3 Model dynamics

Following the above reasoning, the dynamical model involves only one dependent variable: a quantity, R , which represents the absolute risk of burglary, taking into account all pertinent factors. For concreteness, R is defined as the instantaneous rate of burglary, per unit time, at a given location; for each link, e_i , of the graph G , $R_i(t)$ represents the rate at time t . For reasons of generality, time is measured here in non-dimensional units; for a practical implementation these could be specified, and parameters calibrated accordingly. Broadly, though, the time-scale can be assumed to correspond to that over which repeat victimisation effects are evident; an approximate lifetime of two weeks has been observed (Johnson *et al.*, 1997).

Two attributes are associated with each link, e_i : a fundamental *intrinsic attractiveness*, I_i , and a *centrality* measure, C_i . I_i summarises all static factors which influence the decision to commit a burglary at one of the houses on e_i , as perceived by a generic offender when he/she encounters them. These include, for example, the level of physical security and the apparent affluence of the properties; crucially not, though, the level of exposure to potential offenders. In terms of the elements required for the occurrence of crime (Cohen & Felson, 1979), the attractiveness represents the suitability of the target; not the presence of an offender.

The quantity C_i , on the other hand, is a measure of how central link e_i is, from the perspective of travel through the network. Its purpose is to represent imbalances in segment use due to purposeful travel through the network, which were shown in Chapter 3 to be associated with burglary risk. While ‘central’ is not defined in specific terms here (and, notionally, any value could be used), the intention is that it represents a defined network metric, with betweenness being a natural choice. Betweenness was the metric for which a relationship was demonstrated in Chapter 3, and the model is developed with this interpretation of C_i in mind.

As in the model of Short *et al.* (2008), the burglary rate on a link, e_i , is taken

to be composed of two parts: a static component, S_i , and dynamic component, $Q_i(t)$:

$$R_i(t) = S_i + Q_i(t) \quad (5.14)$$

The static rate, which relates to those crimes which are not attributable to boost effects, is assumed to be dominated by opportunistic burglary, and therefore determined by a combination of activity patterns and the intrinsic attractiveness of targets. The calculation of S_i therefore requires an estimate of the activity, W_i , on a link e_i , and the probability, p_i , that an opportunist offender would victimise the link in a given time unit, were he or she to be present.

The activity on each link is defined as the sum of its centrality, C_i , and a constant factor, D , representing the *background activity*; that is, activity which is not determined by the network structure. Activity is therefore given by the expression

$$W_i = D + C_i. \quad (5.15)$$

Mathematically, the effect of D is to mediate between extreme cases: that where activity is homogeneous, and that where it is determined entirely by street network centrality.

The probability of offending, p_i , is modelled as a Poisson process, the rate of which is determined by the intrinsic attractiveness, I_i , defined previously. This implies that the occurrence of burglaries is determined by a series of independent offender decisions, each based on the underlying attractiveness of properties. Accordingly, the probability of an offence within a given time period $[t, t + \delta t)$ is

$$p_i = 1 - e^{-I_i \delta t}. \quad (5.16)$$

In more concrete terms, this implies that, as the attractiveness value I_i increases, the exponential term $e^{-I_i \delta t}$ will tend towards zero and the probability of an offence

therefore approaches 1. Overall, the static component of offending is given by

$$S_i = W_i p_i = (D + C_i) p_i. \quad (5.17)$$

Turning now to the dynamic component of risk, $Q_i(t)$, this term is intended to encapsulate ‘boost’ effects: those acting locally and at short time scales, as a result of recent offending. Its value, for a given link, will again be composed of several parts: that which is due to ‘new’ offending; residual effects from earlier boosts; and boost effects which have spread from neighbouring locations. The first component - new offending - is modelled as increasing by a constant factor, Γ , for every offence which takes place on the link. Taking expected values, this contributes a factor of $\Gamma R_i(t)$ to the rate of change of $Q_i(t)$. Boost effects due to previous incidents persist, but their diminishing influence is reflected in the inclusion of a decay term, $\omega Q_i(t)$, where ω is again a parameter.

Finally, the spread of risk across space is represented as a simple diffusion process between neighbouring links. Specifically, each link e_i is coupled to all other links with which it is co-incident (*i.e.*, other street segments with which it shares a junction) via independent simple diffusion processes, with rate η . The general form for the evolution of $Q_i(t)$ is therefore:

$$\frac{dQ_i}{dt} = \Gamma R_i - \omega Q_i + \eta \sum_{j \sim i} (Q_j - Q_i), \quad (5.18)$$

where \sim indicates the relation ‘ e_i and e_j share a common vertex’. The latter term simply encodes the principle that risk will diffuse from higher- to lower-risk segments: if segment e_i has a neighbour e_j which is at higher risk (*i.e.* $Q_j > Q_i$), the term $(Q_j - Q_i)$ will be positive and drive an increase in Q_i . The summation simply accumulates this effect over all neighbours.

Naturally, R_i can be re-written as $(S_i + Q_i)$, and S_i in turn as $(D + C_i)p_i$, to

give a single dynamical equation for Q_i :

$$\frac{dQ_i}{dt} = \Gamma((D + C_i)p_i + Q_i) - \omega Q_i + \eta \sum_{j \sim i} (Q_j - Q_i). \quad (5.19)$$

5.4 Model analysis

5.4.1 Initial numerical results

In order to demonstrate some of the basic behavioural properties of the model, several numerical simulations are now presented. These use the example street network introduced in Section 5.3.1 to illustrate two aspects of model behaviour: segment-to-segment diffusion, and the influence of centrality. All simulations were performed using a discrete numerical scheme, with time-step δt equal to 0.01.

One of the primary motivations for situating the model on a network is to constrain the diffusion of risk, ensuring that its spread conforms to the shape of the urban backcloth. The principle is most clearly demonstrated by observing the response of the system to individual perturbations of particular links, *i.e.* the spread of risk in the aftermath of hypothetical burglary incidents. To do this, simulations were performed in which the effect of ongoing offending was removed from the model (by setting $\Gamma = 0$) and the system initialised by setting the initial condition of Q to be uniformly 0, except for on one particular link, for which it was set to Ψ . Effectively, then, the simulations considered perturbations to particular dimensions of the diffusion-decay model given by

$$\frac{dQ_i}{dt} = -\omega Q_i + \eta \sum_{j \sim i} (Q_j - Q_i). \quad (5.20)$$

Figure 5.4 shows diffusion patterns corresponding to perturbations in two different locations on the network, labelled e_1 and e_2 . The two streets differ significantly in their position in the network - e_1 is a side-road, whereas e_2 is well-connected and lies on a main route - and the patterns are conspicuously different. When e_1 is victimised

in Figure 5.4a, the spread is localised and directed only to the nearby thoroughfare, whereas risk spreads much more diversely when e_2 is victimised in Figure 5.4b. It is also notable that, in each case, there is negligible spread of risk from e_1 to e_2 , even though they are close in a Euclidean sense. This is, of course, because they are well-separated in terms of network distance; however, such behaviour is contrary to what is predicted by models which do not incorporate the street network.

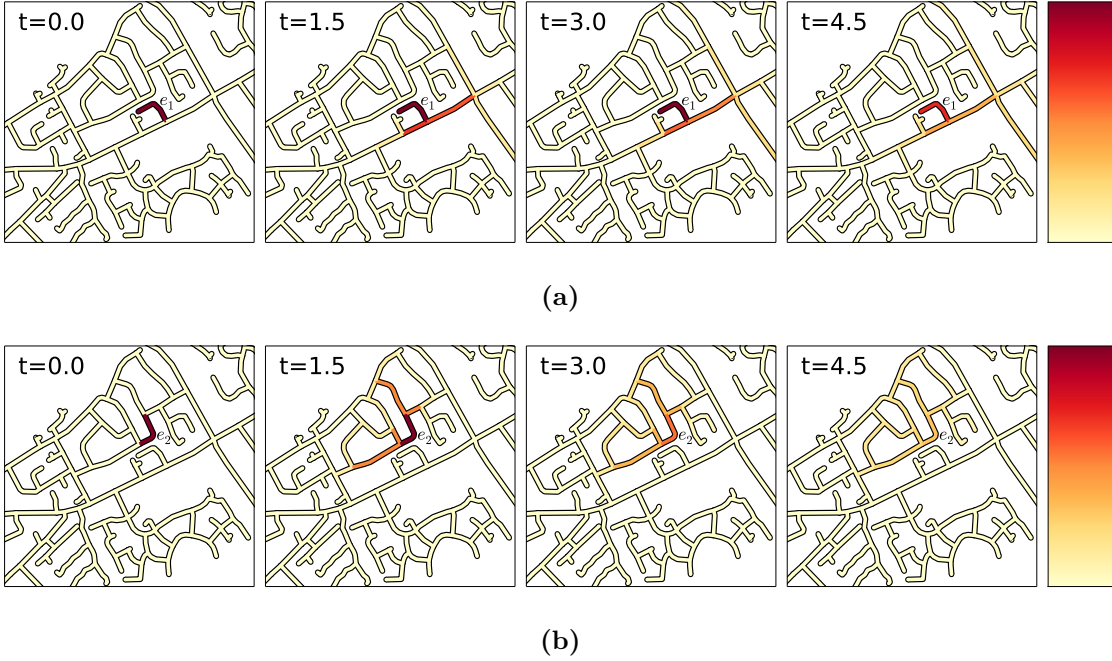


Figure 5.4: Examples of the diffusion of crime risk on a typical section of street network, where the system was perturbed by raising the dynamic burglary risk, Q_i , by Ψ on one segment (as would occur in the event of a burglary). Each set of frames (a) and (b) shows the network, coloured according to Q , at the point of perturbation and at three later times. The parameter values used here are $\eta = 0.25$, $\Psi = 0.4$ and $\omega = 0.1$, and the value of δt used in the simulations is 0.01.

Although these simulations are not instructive in terms of providing an explanation for hot-spot formation, they do demonstrate the potential use of this model as a predictive tool: given a set of recent offences, a quantitative indication of the resulting change in risk distribution can be found.

When the boost term of the equation is re-introduced (by taking $\Gamma \neq 0$), the output of the model instead represents a long-term average of the likely distribution of risk.

In this case, offending occurs and feeds back into the system; rather than being introduced as an exogenous initial condition, burglary activity originates within the model itself.

The primary issue of interest in this case is to examine how the distribution of risk is affected by features of the underlying spatial domain. In particular, since centrality is a key theoretical aspect of the model, the question arises of how its influence is manifested in model results.

A natural example to consider is that where the centrality of each link, e_i , is taken to be equal to its betweenness centrality, which will be denoted by C_i^B . To recap, this is calculated, for a given link e_i , as

$$C_i^B = \sum_{\substack{v_p, v_q \in V, \\ v_p \sim v_q}} \frac{\sigma_{pq}(e_i)}{\sigma_{pq}} \quad (5.21)$$

where σ_{pq} is defined as the total number of shortest paths between v_p and v_q , $\sigma_{pq}(e_i)$ is the total number of shortest paths between v_p and v_q **which pass through the link** e_i , and \sim represents the relation ‘there exists a path between v_p and v_q ’. Figure 5.5 shows the example street network with links coloured according to C^B (which is calculated using an extended network to mitigate edge effects).

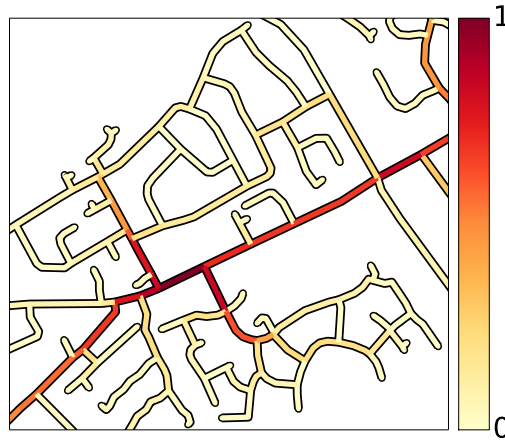


Figure 5.5: Example street network, with street segments coloured according to betweenness. Values are normalised according to the maximum value across the segments shown.

In the corresponding simulation, the intrinsic attractiveness, I_i , was taken to be uniformly equal to 1 across all links, and D was set to 0. In addition, the system was initialised with a uniform value of 0 for Q_i , and the other parameters were set as $\eta = 0.25$, $\Gamma = 0.1$ and $\omega = 0.2$. With the system configured in this way, it was seen numerically that the system tends towards an equilibrium (*i.e.* the value for each link reaches a steady state), with the final state shown in the final frame of Figure 5.6.



Figure 5.6: The evolution of the system towards equilibrium when the static risk is taken to be proportional to betweenness. The parameter values are $\eta = 0.25$, $\Gamma = 0.1$ and $\omega = 0.2$, and the value of δt used in the simulations is 0.01.

Although comparison with Figure 5.5 shows that, as expected, links with high betweenness tend to take large equilibrium values of Q , it is also apparent that the relationship between the two quantities is not a direct one. This can be seen more explicitly in Figure 5.7, in which the betweenness value, C^B , of each segment is plotted against its long-term equilibrium value for Q . In particular, no values of Q lie below 0.14, and some low- C^B segments take values of Q which are much higher than would be the case if the two quantities were proportional. Examples of the latter can be seen clearly in Figures 5.5 and 5.6: segments which have low C^B but are adjacent to those with high C^B take disproportional values of Q .

Behaviour of this form suggests that the diffusive process adds structure to the crime patterns, even in steady state, over and above that which would be predicted purely on the basis of categorising streets by betweenness. Indeed, it suggests that the risk profiles of street segments must be considered explicitly in the context of

those around them, and that the overall risk surface depends on the structure of the urban space as a whole. This will be explored in the following sections, as will the effect of parameter variation on the patterns seen in Figure 5.7.

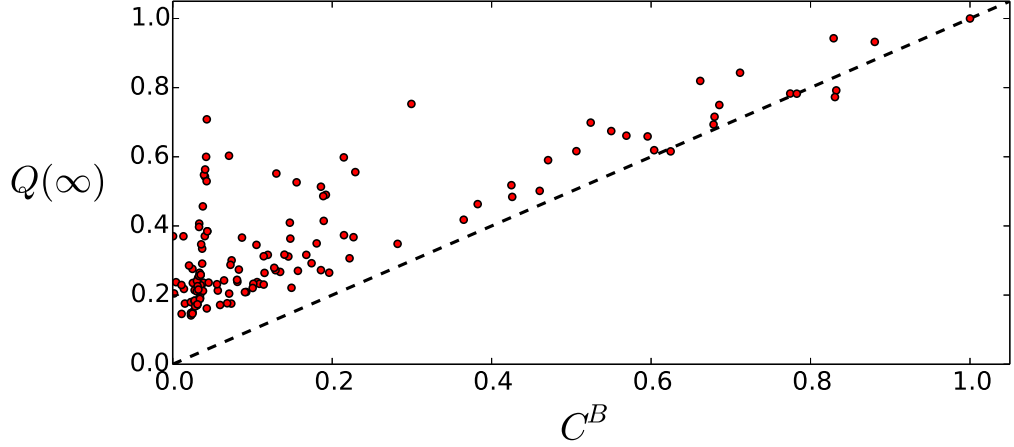


Figure 5.7: Scatter plot of betweenness values, C^B , against long-time equilibrium values of Q , denoted $Q(\infty)$, for the simulation depicted in Figure 5.6. Both values are normalised by their maximum value across the network shown. The dashed line represents equality, *i.e.* the relationship which would be expected if the values were directly proportional.

5.4.2 Algebraic analysis

The numerical results of Section 5.4.1 illustrate that, as expected, a complex relationship exists between the dynamical behaviour of the model and structure of the network on which it is situated. In fact, this relationship can be explicitly characterised in terms of quantities derived from the network, and a closed form for the dynamical evolution also found.

The first step is to introduce a matrix, \mathbf{A}' , which encodes the co-incidence (*i.e.* neighbourhood) of links. $\mathbf{A}' = (a'_{ij})$ is an $m \times m$ matrix, the entries of which are given by

$$a'_{ij} = \begin{cases} 1 & \text{if } e_i, e_j \in E \text{ share a common vertex} \\ 0 & \text{otherwise.} \end{cases} \quad (5.22)$$

In fact, \mathbf{A}' is itself the adjacency matrix for a related graph object: the *line graph* of G , denoted G' . For a given graph, its line graph is found by interchanging the

role of vertices and links: each link is represented by a vertex, and two vertices are linked in G' if the links which they represent are co-incident in G (a full definition and summary are provided by Diestel, 2010). The construction of an example line graph is shown in Figure 5.8 Essentially, G' is the adjacency graph of the links of G , and its adjacency matrix is \mathbf{A}' . Prime notation will be used henceforth to indicate quantities or objects relating to G' .

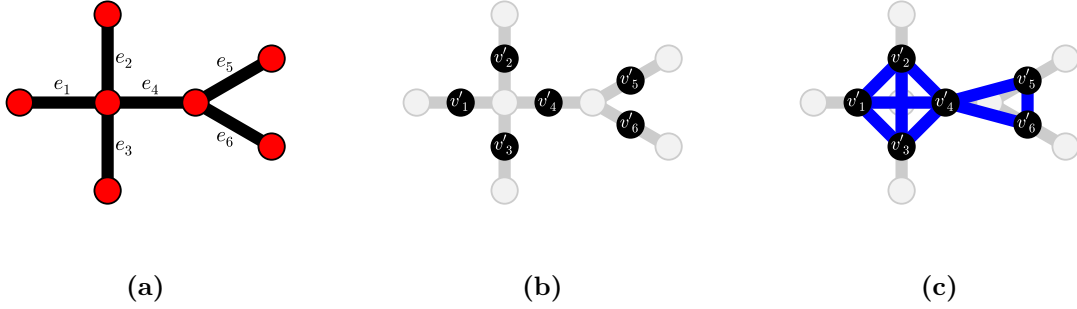


Figure 5.8: The construction of a line graph: a) the original graph G , with links e_i indicated; b) the vertices of the line graph G' , with one vertex, v'_i representing each link of the original graph; and c) the full line graph G' , with two vertices connected if their corresponding links are co-incident in G .

The model equation (5.19) can be re-written in terms of the entries of \mathbf{A}' thus:

$$\frac{dQ_i}{dt} = \Gamma(S_i + Q_i) - \omega Q_i + \eta \sum_j a'_{ij}(Q_j - Q_i), \quad (5.23)$$

where $S_i = (D + C_i)p_i$ has been used for brevity.

A new term, k'_i , is then introduced, representing the ‘degree’ of link e_i (*i.e.* the number of other links with which e_i is co-incident). Noting that k'_i is equal to the number of j for which $a'_{ij} = 1$, (5.23) can be re-arranged and re-written in neater form

as follows:

$$\frac{dQ_i}{dt} = \Gamma S_i + (\Gamma - \omega)Q_i + \eta \sum_j a'_{ij}(Q_j - Q_i) \quad (5.24)$$

$$= \Gamma S_i + (\Gamma - \omega)Q_i + \eta \sum_j a'_{ij}Q_j - \eta k'_i Q_i \quad (5.25)$$

$$= \Gamma S_i + (\Gamma - \omega)Q_i + \eta \sum_j (a'_{ij} - \delta_{ij}k'_i)Q_j \quad (5.26)$$

where δ_{ij} is the Kronecker delta.

A further matrix, $\mathbf{D}' = (d'_{ij})$, is then defined as the diagonal matrix with segment degrees along its diagonal:

$$d'_{ij} = \begin{cases} k'_i & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \quad (5.27)$$

This allows the system represented by (5.26) to be written in vectorised form. Using \mathbf{Q} and \mathbf{S} to represent $(Q_i)_{i=0}^m$ and $(S_i)_{i=0}^m$ respectively, the full system can be written in terms of \mathbf{D}' and \mathbf{A}' as:

$$\frac{d\mathbf{Q}}{dt} = \Gamma \mathbf{S} + (\Gamma - \omega)\mathbf{Q} + \eta(\mathbf{A}' - \mathbf{D}')\mathbf{Q}. \quad (5.28)$$

The diffusive aspect of the model is therefore determined entirely by the bracketed term $(\mathbf{A}' - \mathbf{D}')$. Recalling that \mathbf{A}' and \mathbf{D}' are matrices derived from the line graph G' , this term is in fact the negative of the *graph Laplacian* of G' , defined as $\mathbf{L}' = \mathbf{D}' - \mathbf{A}'$. This is a well-studied object in graph theory (see Section 5.4.3), and substituting it in (5.28) yields

$$\frac{d\mathbf{Q}}{dt} + (\omega - \Gamma)\mathbf{Q} + \eta \mathbf{L}'\mathbf{Q} - \Gamma \mathbf{S} = 0. \quad (5.29)$$

The key step in proceeding from here is to note that the vector \mathbf{Q} can be written as a linear combination of the eigenvectors of the Laplacian; this provides the fundamental link between topology and dynamics. Denoting these eigenvectors as \mathbf{w}_h ,

the vector \mathbf{Q} can be written as

$$\mathbf{Q}(t) = \sum_h q_h(t) \mathbf{w}_h, \quad (5.30)$$

where the coefficients $q_h(t)$ are dynamic. The same can then be done for the static risk \mathbf{S} :

$$\mathbf{S} = \sum_h s_h \mathbf{w}_h, \quad (5.31)$$

where the coefficients s_h are static. These can then be substituted into (5.29), to give

$$\sum_h \left(\frac{dq_h}{dt} + (\omega - \Gamma)q_h + \eta \mathbf{L}' q_h - \Gamma s_h \right) \mathbf{w}_h = 0. \quad (5.32)$$

This equation is in exactly the form for which the standard eigenvector relationship can be applied. Using $\mathbf{L}' \mathbf{w}_h = \lambda_h \mathbf{w}_h$, where λ_h is the eigenvalue corresponding to \mathbf{w}_h , (5.32) becomes

$$\sum_h \left(\frac{dq_h}{dt} + (\omega - \Gamma)q_h + \eta \lambda_h q_h - \Gamma s_h \right) \mathbf{w}_h = 0. \quad (5.33)$$

Now, since the Laplacian is symmetric, its eigenvectors are orthogonal and each term of the summation will therefore hold independently; that is

$$\frac{dq_h}{dt} + (\omega - \Gamma)q_h + \eta \lambda_h q_h - \Gamma s_h = 0, \quad (5.34)$$

for all h . Two distinct cases must be considered when solving this differential equation.

Case 1: $\omega - \Gamma + \eta \lambda_h \neq 0$. In this case, equation (5.34) can be solved via standard methods to give, for all such h

$$q_h(t) = \frac{\Gamma s_h}{\omega - \Gamma + \eta \lambda_h} + \left(q_h(0) - \frac{\Gamma s_h}{\omega - \Gamma + \eta \lambda_h} \right) e^{-t(\omega - \Gamma + \eta \lambda_h)}. \quad (5.35)$$

Case 2: $\omega - \Gamma + \eta\lambda_h = 0$. In this case, q_h no longer appears in equation (5.34), and it reduces to the simple form

$$\frac{dq_h}{dt} - \Gamma s_h = 0. \quad (5.36)$$

The solution to this, corresponding to simple linear growth with time, is given by

$$q_h(t) = q_h(0) + \Gamma s_h t. \quad (5.37)$$

These solutions, when taken together and substituted into (5.30), determine the state of the system for all time t :

$$\begin{aligned} \mathbf{Q}(t) = & \sum_{\substack{h: \\ \omega - \Gamma + \eta\lambda_h \neq 0}} \left[\frac{\Gamma s_h}{\omega - \Gamma + \eta\lambda_h} + \left(q_h(0) - \frac{\Gamma s_h}{\omega - \Gamma + \eta\lambda_h} \right) e^{-t(\omega - \Gamma + \eta\lambda_h)} \right] \mathbf{w}_h \\ & + \sum_{\substack{h: \\ \omega - \Gamma + \eta\lambda_h = 0}} \left[q_h(0) + \Gamma s_h t \right] \mathbf{w}_h. \end{aligned} \quad (5.38)$$

At all times, therefore, \mathbf{Q} can be understood as the superposition of m components; one for each eigenvector \mathbf{w}_h . These components can be interpreted as ‘modes’ of risk - they correspond to vectors over the network - and the magnitude of each is given by an individual expression in t . Since the influence of network structure is captured entirely by the eigenvalues and eigenvectors of \mathbf{L}' , it is necessary to examine the properties of the Laplacian in order to characterise the dynamics.

5.4.3 Properties of the Laplacian

The study of the Laplacian and its properties lies within the more general field of ‘spectral graph theory’, which is concerned with the purely mathematical properties of abstract graphs. A full review is provided by Chung (1997), and the volume by van Mieghem (2012) also gives many key results, with additional emphasis on their application to complex networks (*i.e.* large, real-world graphs). Much of the theory concerning the Laplacian relates to the characterisation of its eigenvalues and their relationship to other graph properties, such as the adjacency matrix. In particular,

a number of theorems establish upper and lower bounds for the eigenvalues in terms of, for example, vertex degrees.

While these results are powerful, their utility in the present context is somewhat limited. As remarked previously, many classical graph properties (such as vertex degree) are relatively uninformative for street networks (and their line graphs), since physical constraints mean their values fall in a very narrow range; many results, therefore, offer little insight into the distribution of eigenvalues. In addition, the exact magnitude of eigenvalues is of relatively little interest in terms of the qualitative behaviour of the model. Nevertheless, some basic results allow immediate progress to be made. Although, for brevity, the results are expressed in terms of the notation introduced for the model above, they apply to any generic graph G .

To re-cap, the Laplacian, $\mathbf{L} = (l_{ij})$, for a graph, G , is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$ (where \mathbf{D} is the diagonal degree matrix and \mathbf{A} the adjacency matrix), so that

$$l_{ij} = \begin{cases} k_i & \text{if } i = j \\ -1 & \text{if } v_i \text{ and } v_j \text{ are adjacent} \\ 0 & \text{otherwise.} \end{cases} \quad (5.39)$$

For undirected graphs (as considered here), \mathbf{L} is symmetric and positive semi-definite. By standard linear algebra, this implies that its eigenvalues λ_h are real and non-negative. For convenience, they can therefore be labelled in ascending order, so that $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.

It is possible to go further than this by making the simple observation that the row-sums of \mathbf{L} are always equal to zero. For a given row i , the number of -1 entries is exactly equal to the number of neighbours v_i has in the graph, which itself is equal to k_i . Multiplying \mathbf{L} by the constant vector $\mathbf{1} = (1, 1, \dots, 1)^T$ therefore gives, in the

i th element of the result:

$$\sum_j l_{ij} \times 1 = \sum_j (\delta_{ij} k_i - a_{ij}) = k_i - \sum_j a_{ij} = k_i - k_i = 0. \quad (5.40)$$

Thus the vector $\mathbf{1}$ is always an eigenvector for \mathbf{L} , with eigenvalue 0. It is therefore the case that $\lambda_1 = 0$ and so

$$0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n. \quad (5.41)$$

In fact, it can be further shown that the multiplicity with which 0 occurs as an eigenvalue is exactly equal to the number of connected components of the graph (see van Mieghem, 2012). For simplicity, and because it is usually the case for street networks, it is assumed here that all graphs under consideration are connected, and so $\lambda_2 \neq 0$.

A further corollary is that, for any eigenvector other than $\mathbf{1}$, the sum of the elements of the vector must be 0. Since all eigenvectors are orthogonal, the dot product of any two must equal 0; in particular, $\mathbf{1}^T \mathbf{w}_h = 0$ for all $\mathbf{w}_h \neq \mathbf{1}$. Since $\mathbf{1}^T \mathbf{w}_h$ is simply the sum of the elements of \mathbf{w}_h , the result is immediate.

Aside from those presented above, bounds for eigenvalues are of little relevance in the present work. Nevertheless, the following bounds are well-known (Mohar, 1991) and can be instructive at certain points (k_{min} and k_{max} are, respectively, the minimum and maximum degree of the graph):

$$\lambda_2 \leq \frac{n}{n-1} k_{min} \leq \frac{n}{n-1} k_{max} \leq \lambda_n \leq 2k_{max} \quad (5.42)$$

Several more spectral results and an outline of their possible applications can be found in the chapter by Mohar (1997).

5.4.4 Implications for model stability

Several of the above observations have implications for the model of burglary. In particular, they allow insight to be gained concerning the long-term behaviour of the model: its stability, and the existence and form of equilibria as $t \rightarrow \infty$.

It has been shown that the component of the model corresponding to dynamic burglary risk \mathbf{Q} is given by

$$\begin{aligned} \mathbf{Q}(t) &= \sum_h q_h(t) \mathbf{w}_h \\ &= \sum_{\substack{h: \\ \omega - \Gamma + \eta\lambda_h \neq 0}} \left[\frac{\Gamma s_h}{\omega - \Gamma + \eta\lambda_h} + \left(q_h(0) - \frac{\Gamma s_h}{\omega - \Gamma + \eta\lambda_h} \right) e^{-t(\omega - \Gamma + \eta\lambda_h)} \right] \mathbf{w}_h \\ &\quad + \sum_{\substack{h: \\ \omega - \Gamma + \eta\lambda_h = 0}} \left[q_h(0) + \Gamma s_h t \right] \mathbf{w}_h. \end{aligned} \quad (5.43)$$

The dependence on t therefore takes two forms: exponential terms in the first series, and linear terms in the second. For both of these, there are simple conditions under which the system will not simply diverge towards infinity: for the exponential terms, the coefficients remain bounded if and only if

$$\omega - \Gamma + \eta\lambda_h \geq 0, \quad (5.44)$$

and, similarly, the linear terms remain bounded if and only if

$$\Gamma s_h = 0. \quad (5.45)$$

It is possible to examine these criteria, and the long-term behaviour of the model, under different configurations of the parameters Γ and ω . In doing so, the key observation concerns the eigenvalues appearing in the equation: by Section 5.4.3, the eigenvalues can be ordered $0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_m$ (since \mathbf{L}' , from which they are derived, is an $m \times m$ matrix, it has m eigenvalues). It will be assumed here that the eigenvalues λ_h and their corresponding eigenvectors \mathbf{w}_h are labelled according

to this order. It is also known that, with this labelling, \mathbf{w}_1 is equal to the constant vector $\mathbf{1}$.

Case 1: $\omega < \Gamma$. If $\omega < \Gamma$, then $\omega - \Gamma + \eta\lambda_1 = \omega - \Gamma < 0$. The coefficient of \mathbf{w}_1 in equation (5.43) therefore diverges to infinity, and so \mathbf{Q} is unstable in this case.

Case 2: $\omega = \Gamma$. In this case, the second summation of equation (5.43) contains only one term - the coefficient of \mathbf{w}_1 - since:

$$\begin{aligned}\omega - \Gamma + \eta\lambda_h &= 0 & h = 1 \\ \omega - \Gamma + \eta\lambda_h &\neq 0 & 2 \leq h \leq m\end{aligned}$$

In order to remain bounded, therefore, it is necessary that $\Gamma s_1 = 0$; however, this represents a trivial case. If $\Gamma = 0$, the coefficients of \mathbf{w}_h for $2 \leq h \leq m$ all tend to 0 (since all exponential terms are decaying), so that the eventual equilibrium solution, denoted $\overline{\mathbf{Q}}$, is the trivial homogeneous case:

$$\overline{\mathbf{Q}} = q_1(0)\mathbf{w}_1 = q_1(0)\mathbf{1}. \quad (5.46)$$

The possibility that $s_1 = 0$ requires slightly deeper consideration. Recalling that the values s_h are defined as the solutions to

$$\mathbf{S} = \sum_{h=1}^m s_h \mathbf{w}_h = s_1 \mathbf{1} + \sum_{h=2}^m s_h \mathbf{w}_h, \quad (5.47)$$

the key observation is that, for $2 \leq h \leq m$, the sum of the elements of \mathbf{w}_h is 0. This means that s_1 is equal to the mean of the elements of \mathbf{S} ; since all elements of \mathbf{S} are non-negative, a value of $s_1 = 0$ implies that \mathbf{S} is identically zero. This, in turn, implies that $s_h = 0$ for all values of h . As with Case 1, this leads to a trivial homogeneous equilibrium:

$$\overline{\mathbf{Q}} = q_1(0)\mathbf{w}_1 = q_1(0)\mathbf{1}. \quad (5.48)$$

Case 3: $\omega > \Gamma$. It is this case which gives rise to the most interesting behaviour. Because all eigenvalues are non-negative, $(\omega - \Gamma + \eta\lambda_h)$ is positive for all h , and so equation (5.43) contains only coefficients of the exponential type. In addition, this positivity ensures that all exponential terms decay to 0, so that

$$\lim_{t \rightarrow \infty} q_h(t) = \frac{\Gamma s_h}{\omega - \Gamma + \eta\lambda_h} \quad \forall h. \quad (5.49)$$

In this case, therefore, the system converges to an equilibrium, the value of which is given by

$$\bar{\mathbf{Q}} = \sum_{h=1}^m \frac{\Gamma s_h}{\omega - \Gamma + \eta\lambda_h} \mathbf{w}_h. \quad (5.50)$$

The known values of $\lambda_1 = 0$ and $\mathbf{w}_1 = \mathbf{1}$ can then be substituted to give the more explicit form

$$\bar{\mathbf{Q}} = \frac{\Gamma s_1}{\omega - \Gamma} \mathbf{1} + \sum_{h=2}^m \frac{\Gamma s_h}{\omega - \Gamma + \eta\lambda_h} \mathbf{w}_h. \quad (5.51)$$

Again, the fact that the eigenvectors $\mathbf{w}_2, \dots, \mathbf{w}_m$ all sum to zero can be used to characterise this equilibrium value further. Since these eigenvectors do not contribute to the element-wise mean of $\bar{\mathbf{Q}}$, the mean value of \bar{Q}_i across all elements must therefore be given by $\frac{\Gamma s_1}{\omega - \Gamma}$. The effect of the eigenvectors $\mathbf{w}_2, \dots, \mathbf{w}_m$ is therefore to introduce variation in the value of \bar{Q}_i on individual segments, around this central value.

Since it is the only case in which $\mathbf{Q}(t)$ remains bounded, and for which the equilibrium state is not homogeneous, it will be assumed in the remainder of the work that $\omega > \Gamma$.

5.4.5 Relationship between static risk and equilibrium

The numerical results shown in Figures 5.6 and 5.7 demonstrated the existence of a non-linear relationship between the static risk, S_i , on each link and the corresponding equilibrium value of dynamic risk, \bar{Q}_i . By comparing equation (5.51) with the eigenvector decomposition of \mathbf{S} , the reason for this becomes clear. Since

$$\mathbf{S} = s_1 \mathbf{1} + \sum_{h=2}^m s_h \mathbf{w}_h, \quad (5.52)$$

it is seen that $\overline{\mathbf{Q}}$ is related to \mathbf{S} by a multiplicative factor for each term in the summation. Rather than simply being multiplied by a constant factor $\frac{\Gamma}{\omega-\Gamma}$, though, each coefficient in (5.51) also includes the additional term $\eta\lambda_h$ on the denominator, which is strictly positive for $2 \leq h \leq m$.

The effect of this is that, in moving from the static risk to the equilibrium value of \mathbf{Q} , the contribution of each eigenvector \mathbf{w}_h to variation around the mean value $\frac{\Gamma s_1}{\omega-\Gamma}$ is damped. Furthermore, the size of this damping is determined by the magnitude of the corresponding eigenvalue: the higher the eigenvalue, the lower the value of $\frac{\Gamma s_h}{\omega-\Gamma+\eta\lambda_h}$ and so the more the influence of that mode is diminished.

In qualitative terms, the equilibrium distribution of risk reached by the model is a ‘smoothed’ version of the static risk distribution, in which the variation corresponding to higher-eigenvalue modes is diminished most. To understand how such eigenvectors relate to the spatial structure of the network, it is instructive to examine the eigenvalues and eigenvectors of example networks.

5.4.6 Spectra of street networks

Since the real-world example network used in this chapter is somewhat intricate, a stylised hypothetical example will be considered first, in order to afford a general understanding. Figure 5.9 shows a simple network of only 18 links, constructed to reflect loosely the tree-like character of a hierarchical residential street network.

In Figure 5.9a, its links are coloured according to betweenness, with the highest values occurring on the most central streets, as expected. In Figure 5.9b, the colouring shows the equilibrium values of Q_i for a simulation in which $\eta = 0.1$, $\Gamma = 0.1$ and $\omega = 0.2$, and static risk S_i set to be proportional to betweenness. In both cases the values have been normalised for comparison, and it is clear that the variation in \overline{Q}_i across links is notably less than that for S_i .

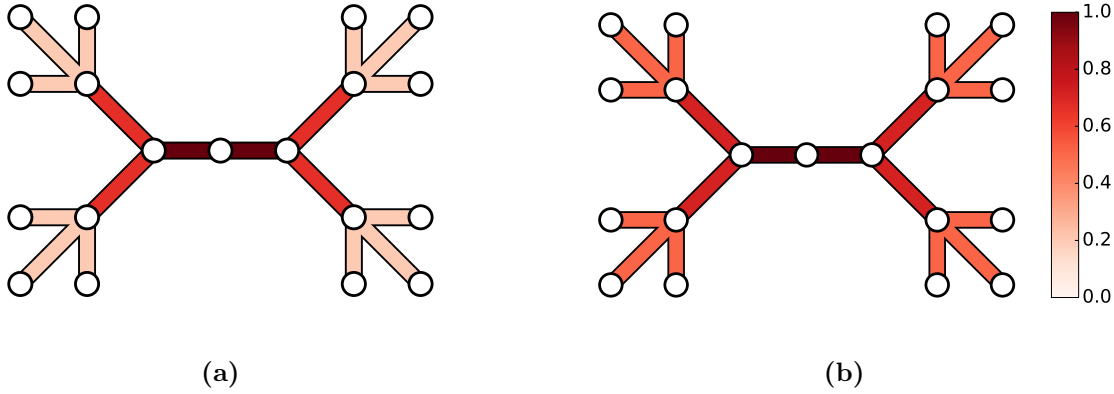


Figure 5.9: A stylised hypothetical street network, with links coloured according to: a) betweenness; and b) equilibrium values of Q_i for the burglary model with parameter values $\eta = 0.1$, $\Gamma = 0.1$ and $\omega = 0.2$. Colours are normalised in each case according to the maximum value.

The line graph for this network can be computed, as can its Laplacian. The Laplacian has 18 eigenvalue-eigenvector pairs, and each eigenvector has 18 elements: one for each vertex of the line graph (*i.e.* each link of the original graph). Figure 5.10 shows all of these eigenvectors: each panel represents an eigenvalue, and the links of the network are coloured according to the corresponding element of the associated eigenvector. Red and blue indicate positive and negative values respectively, and, for ease of comparison, each eigenvector has been rescaled so that the maximum of its elements' absolute values is equal to 1.

Although no general characterisation of the eigenvectors is immediately apparent, several qualitative observations can be made. Firstly, it appears that, as the eigenvalues increase, there tend to be sharper discontinuities between the eigenvector values for neighbouring links. The vector associated with λ_2 , for example, transitions smoothly between two clear 'halves' of the network, and a similar principle is evident for λ_3 and λ_4 . For the relatively large eigenvalues, though, the eigenvectors show considerable localised variation. In a sense, the spatial structure of the network is 'refined' as subsequent eigenvectors accumulate.

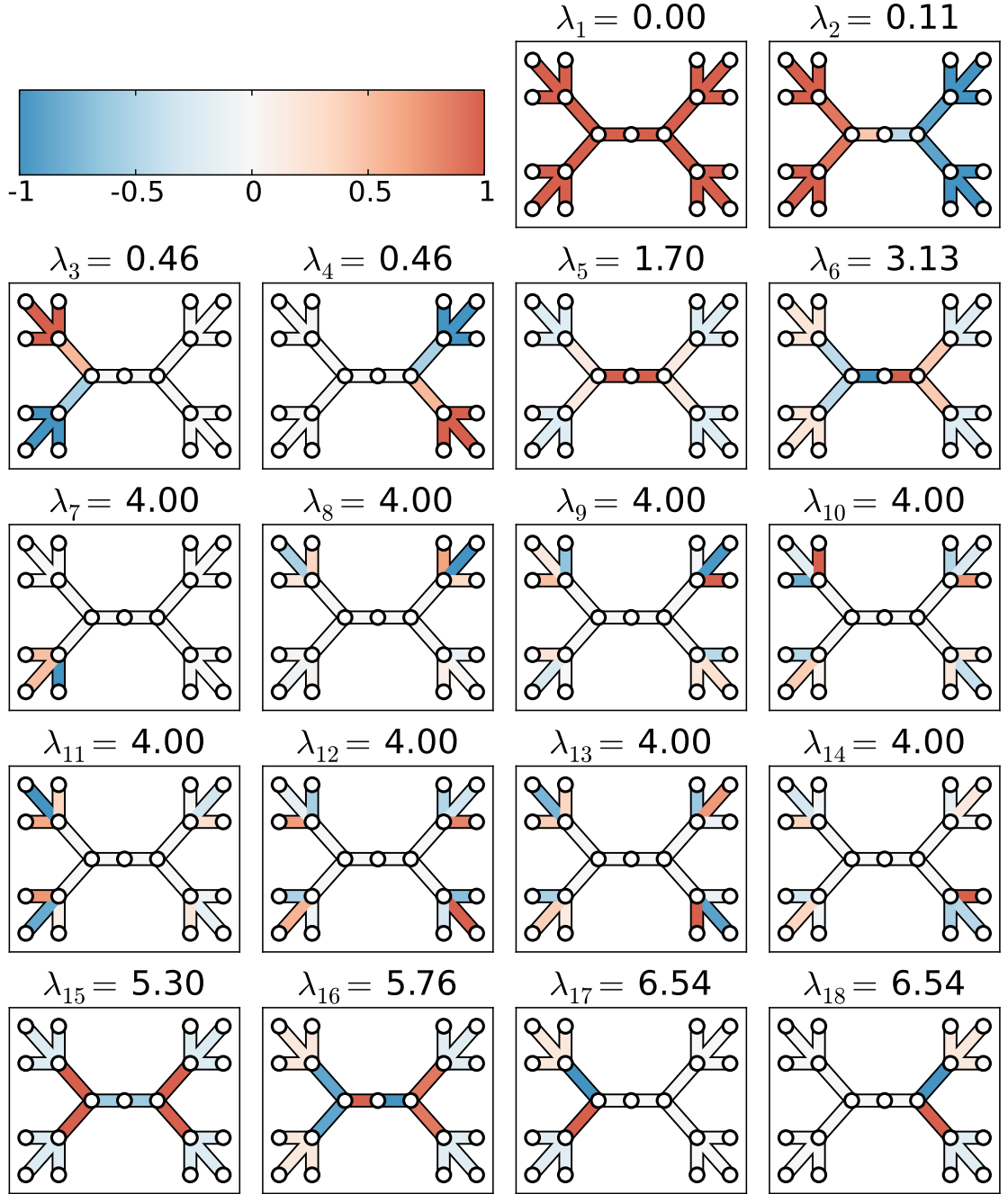


Figure 5.10: The eigenvalues and eigenvectors of the simple example network, ordered with increasing eigenvalue. In each case, the links of the network are coloured according to the (normalised) value of the eigenvector in the corresponding dimension.

Secondly, there is a pattern of symmetry in the eigenvectors associated with repeated eigenvalues; this is apparent for $\lambda = 0.46$, 4.00 and 6.54 . The symmetry can be broadly reconciled with community structures in the network, and is an expected artefact of the clear regularity present in the network.

The third observation concerns how network betweenness might be expressed in terms of these vectors. As seen in Figure 5.9a, the values of betweenness exhibit both left/right and top/bottom symmetry (when the network is arranged in this way), and so eigenvectors which do not have that property would not be expected to feature in its decomposition. The eigenvectors associated with λ_1 , λ_5 and λ_{15} are the only ones which are symmetrical in this way.

This is confirmed when the betweenness vector for the example network is decomposed in terms of the eigenvectors, *i.e.* when \mathbf{S} is taken to be proportional to betweenness (as shown in Figure 5.9a), and the equation

$$\mathbf{S} = \sum_h s_h \mathbf{w}_h \quad (5.53)$$

is solved for s_h . In order that the magnitude of the coefficients can be meaningfully compared, the eigenvectors \mathbf{w}_h are first normalised before solving. Table 5.1 shows the solution for the three eigenvectors for which $s_h \neq 0$, which are the three anticipated.

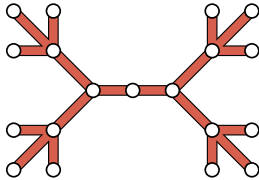
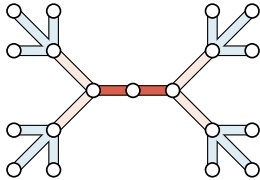
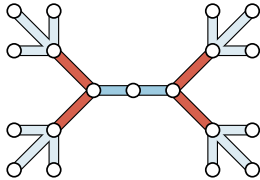
h	1	5	15
λ_h	0.0	1.7	5.3
s_h	149.91	103.94	34.83
\mathbf{w}_h			

Table 5.1: The eigenvectors which appear with non-zero magnitude in the solution of $\mathbf{S} = \sum_h s_h \mathbf{w}_h$, where \mathbf{S} is taken to be proportional to betweenness.

Referring back to equation (5.50), the equilibrium solution in Figure 5.9b can be seen as a superposition of the 3 modes shown in Table 5.1. The smoothness, by comparison with the underlying static risk, is a result of the damped influence of modes \mathbf{w}_5 and \mathbf{w}_{15} . The variation in the latter direction is diminished to a greater extent, due to its larger associated eigenvalue.

5.5 Model dynamics and community structure

An immediate observation when considering the eigenvectors shown in Figure 5.10 is the apparent correspondence with community structure in the network. In the ordering given, the first few eigenvectors appear to identify coarse-scale communities in the network (informally ‘halves’ and ‘quarters’, for example), before this becomes more granular as the associated eigenvalues increase. Indeed, this relationship - between communities and the eigenvalues of the Laplacian - has been the subject of a significant volume of research (see Newman, 2013b), and several formal results have been established.

Furthermore, this relationship also extends to the study of dynamical systems on networks. The effect of community structure can be observed in the behaviour of a number of well-studied examples, the dynamics of which are seen to be unified, to some extent, within communities. Indeed, the converse relationship can also be exploited: the behaviour on a network of certain dynamical systems can be used to infer community structure.

The potential existence of community-level effects in a burglary model is also of real-world relevance. In street networks, communities correspond to notions of neighbourhood and insularity; furthermore, they are meaningful units in terms of policing provision. If discontinuities in burglary risk are aligned with such structures, this suggests that such areas can be regarded as identifiable units of common crime risk, and policed accordingly. The remainder of the chapter will examine such effects in

the proposed burglary model via both analytical and simulation-based methods. In order to motivate this, some background research will first be reviewed.

5.5.1 Previous research

Community identification is one of the central themes of research within network science, and has clear real-world importance in topics such as social networks (see Fortunato, 2010, for a review). In general, the objective of community identification is to partition the vertices of a network into groups, such that connectivity is dense within groups but sparse between them (where ‘connectivity’ typically refers to the number of links). When such a partition can be found, it can be taken to define a natural division, in some sense, of the vertices. Beyond this general definition, however, there are variations in how the problem is posed: what constitutes a ‘good’ division, for example, or whether the number of communities to be found is pre-specified or somehow inferred.

One of the simplest formulations concerns how a graph can be split into two groups, of given size, so that the number of links between them is minimised. Although the pre-specification of group sizes is not realistic in many applications, there are certain circumstances in which it is appropriate. An example is in the design of algorithms in computing science: if vertices represent dependencies between processing tasks to be distributed between two processors of known capacities, and inter-processor communication is slow, the problem corresponds to finding the most efficient division of tasks.

In this case, the method of ‘spectral partitioning’ can be used to find an approximate solution. In this method, an *index vector*, \mathbf{h} , is defined as

$$h_i = \begin{cases} +1 & \text{if vertex } i \text{ belongs to group 1} \\ -1 & \text{if vertex } i \text{ belongs to group 2,} \end{cases} \quad (5.54)$$

which encodes the assignment, and the task is therefore to find a choice of \mathbf{h} which minimises the number of links between groups, defined as R . It can be shown (see Riolo & Newman, 2014) that

$$R = \frac{1}{4} \mathbf{h}^T \mathbf{L} \mathbf{h} \quad (5.55)$$

where \mathbf{L} is the Laplacian of the graph, and this provides the crucial link to spectral theory. If the index vector is written as a sum of the eigenvectors \mathbf{w}_i of \mathbf{L} (ordered, in the usual way, according to increasing eigenvalue), so that $\mathbf{h} = \sum_i a_i \mathbf{w}_i$, R reduces to

$$R = \sum_i a_i^2 \lambda_i. \quad (5.56)$$

The task is then to find the a_i which will minimise R , under the constraint that $\sum_i a_i^2 = N$. Given the properties of the Laplacian ($\lambda_1 = 0$, and all other eigenvalues are strictly positive), this task is simple: to choose a_i so that the majority of weight in the summation is in the early terms. Since the value of a_1 is constrained by the fixed group sizes, the lowest value for R is therefore found by placing as much weight as possible in the a_2 term.

Although choosing \mathbf{h} parallel to \mathbf{w}_2 is not, in general, possible, since the elements of \mathbf{h} must take the values ± 1 , this can be approximated by taking the elements of \mathbf{h} to have the same sign as their counterparts in \mathbf{w}_2 . After some adjustment to ensure that the group sizes are correct, this gives an assignment which has been shown to perform well when applied to real networks (Newman, 2013a).

It is clear, therefore, that the vector \mathbf{w}_2 has a close relationship with two-way partitioning of networks (the correspondence is evident, for example, in Figure 5.10). Clearly, though, division into two groups of known size is of limited utility, and attempts have been made to generalise the approach to a larger number of groups. The work of Riolo & Newman (2014), who examined that case where k groups are required, is a particularly clear example. Working from first principles, they derived a method which also involves the assignment of vertices to groups on the basis of eigenvector values; however, rather than using just \mathbf{w}_2 , the assignment was found

by combining vectors \mathbf{w}_2 to \mathbf{w}_k . This further illustrates the principle that the eigenvectors of the Laplacian define an increasingly granular community structure, when ordered according to eigenvalue magnitude. In fact, it has recently been shown by Newman (2013a) that, at least in the two-group case, other forms of community detection (*i.e.* those which do not simply minimise the number of inter-group links) are, in fact, equivalent to the spectral method.

The eigenvalues of the Laplacian have also been examined in the context of dynamical systems taking place on networks. The term ‘dynamical system’ is defined broadly and refers to any system for which the dynamics on a given vertex are determined by a combination of internal activity and a simple coupling with neighbouring vertices: examples include Kuramoto oscillators (Arenas *et al.*, 2006a), diffusion (Cheng & Shen, 2010) and majority dynamics (Almendral & Díaz-Guilera, 2007). The questions of interest typically concern synchronisation - a topic which is ubiquitous in complexity science - and the stability of synchronised states.

The early work of Barahona & Pecora (2002) identified a fundamental condition for the synchronisability (existence of a stable synchronised state) of a system of oscillators. The crucial condition obtained is that a state is stable if the ‘eigenratio’ $\frac{\lambda_n}{\lambda_2}$ is less than a certain value. Stability is therefore determined by only two eigenvalues of the Laplacian - the first non-zero eigenvalue and the largest - and, since the largest is primarily determined by the maximum degree of the network, the main dependence is on λ_2 . This is a particularly striking result, especially when the known relationship with community structure is taken into account.

Building on that work, Arenas *et al.* (2006b) also considered networks of coupled oscillators, but from the point of view of the path to synchronisation, *i.e.* the dynamical behaviour before the synchronised state was reached. The authors constructed networks with clear hierarchical community structure (4 strong communities, each of which had 4 sub-communities) and performed repeated simulations of oscilla-

tors with random initial states. The pairwise synchronisation between all pairs of oscillators was quantified at each time and converted to a binary value by thresholding (so that any pair were either ‘synchronised’ or ‘not synchronised’ at any time).

The results show that the synchronisation in the network occurs in a sequential, hierarchical fashion: the 16 lowest-level sub-communities synchronise within themselves first, followed by the 4 larger communities, and finally the network as a whole. In addition, the synchronisation phases for each hierarchical level remain stable for relatively long periods: the low-level synchronisation persists for some time, before a fast transition to synchronisation at the higher community level. These ‘stability plateaus’ reflect, in some sense, the stability of synchronised states at the scales defined by the communities.

In addition to this, a notable relationship was observed between the distribution of eigenvalues and the duration of each stability plateau. In particular, it was shown that the gaps in the ordered series of inverted eigenvalues $\frac{1}{\lambda_i}$ correspond to the periods over which the number of synchronised components remains constant (*i.e.* the lifetime of each successive stability plateau). To be precise, the relative gap between $\frac{1}{\lambda_{k+1}}$ and $\frac{1}{\lambda_k}$ corresponds to the relative duration of the period over which k components are synchronised; in general, the relative length of the interval $\left[\frac{1}{\lambda_{k+1}}, \frac{1}{\lambda_k}\right)$ corresponds to the stability of the k -community stage.

5.5.2 Analytical argument for community effects

The reason for the behaviour described above can be seen by considering the analytical solution for systems of this type. After simplification, the coupling in systems of this type reduces to a general form

$$\frac{d\mathbf{X}(t)}{dt} = -r\mathbf{L}\mathbf{X}(t), \quad r > 0 \quad (5.57)$$

where $X(t)$ is some quantity defined at every vertex: this is almost immediate in the case of diffusion-like processes, and arises in the linearisation of oscillator systems.

Following the same reasoning as Section 5.4.2 (and continuing with previous notation for eigenvalues and eigenvectors), the solution to this is given by

$$\mathbf{X}(t) = \sum_i \phi_i(t) \mathbf{w}_i = \sum_i \phi_i(0) e^{-\lambda_i r t} \mathbf{w}_i \quad (5.58)$$

where the $\phi_i(t)$ are the coefficients of each of the eigenvector modes.

As time proceeds, therefore, the mode coefficients $\phi_i(t) = \phi_i(0) e^{-\lambda_i r t}$ decay to zero (for all $i \neq 1$), each at a different rate. Although the time taken to reach zero is infinite in all cases, a lifetime τ_i can be defined as the time taken for $\phi_i(t)$ to fall below a given small threshold ϵ (regarded, for practical purposes, as zero). With ϵ selected in this way, the lifetime for each i is given by

$$\tau_i = \frac{1}{\lambda_i} \times \frac{\log \phi_i(0) - \log \epsilon}{r}. \quad (5.59)$$

The moments τ_i therefore form a series of time intervals $[0, \tau_n), [\tau_n, \tau_{n-1}), \dots, [\tau_3, \tau_2), [\tau_2, \infty)$ which divide the dynamical process into n stages. To be specific, the interval $[\tau_{i+1}, \tau_i)$ can be regarded as stage i of the diffusion process (reversing the temporal ordering for notational convenience, so that stage n is the first to occur). During stage i , coefficients $\phi_j(t)$ have decayed to zero for all $j > i$, and so the solution is given by

$$\mathbf{X}(t) = \sum_{j=0}^i \phi_j(t) \mathbf{w}_j. \quad (5.60)$$

This explains the evolution of community structure: as each stage ends, the variation in the direction of the corresponding vector is nullified, and the vertices are synchronised with respect to that vector. Going further, this reasoning also accounts for the correspondence between gaps in the eigenvalue spectrum and the stability of semi-synchronised states. As seen from (5.59), the times of transitions between stages (*i.e.* the times at which modes disappear) are determined by the inverted eigenvalues, and so persistent states correspond to large gaps in the $\frac{1}{\lambda_i}$ sequence. In this sense, therefore, the stability of an i -community synchronised state corresponds to the length of the interval $\left[\frac{1}{\lambda_{i+1}}, \frac{1}{\lambda_i} \right)$.

This behaviour is illustrated in Figure 5.11, in which a number of stages in the evolution of $\mathbf{X}(t)$ are identified. The system is initialised with random noise and allowed to evolve to a fully-homogeneous state. The three moments τ_2 , τ_3 and τ_4 are shown; although not to scale, their uneven spacing is correct, and corresponds to the relative stability of the final stages. Prior to the elimination of mode \mathbf{w}_4 at τ_4 , the system shows approximate synchronisation within 4 communities. When τ_4 is reached, there is then a rapid transition to 2-community synchronisation (since $\tau_3 \approx \tau_4$, the 3-community state is transient), which remains until the nullification of the final non-constant mode, \mathbf{w}_2 , at τ_2

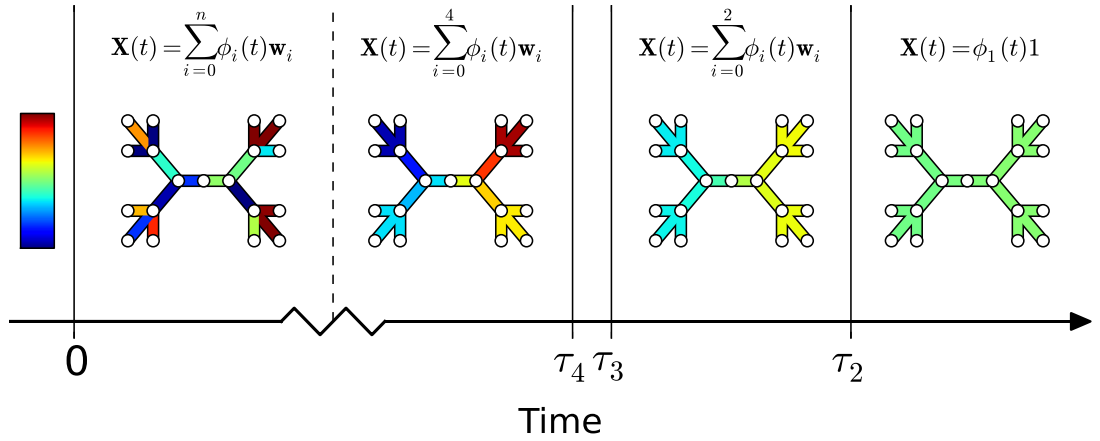


Figure 5.11: The correspondence between intra-community synchronisation and eigenvalues for a simple diffusion process taking place on a network. The moments τ_i of a number of modes are shown (not to scale), and the approximate state of the system during each inter-moment interval is indicated by the colouring of the network. The accompanying formulae show the value of $\mathbf{X}(t)$ in terms of those modes whose magnitudes have not fallen below ϵ .

One final observation concerns the particular status of the eigenvalue λ_2 . It was noted previously that larger values of λ_2 correspond to greater dynamical stability (Barahona & Pecora, 2002), and it is now clear that, when synchronisation does occur, smaller values of λ_2 have the effect of extending the period until synchronisation is achieved (since it is the last non-constant mode, and its lifetime is proportional to $\frac{1}{\lambda_2}$). In fact, a further link can be observed by recalling that the multiplicity of 0 as an eigenvalue gives the number of connected components of a graph: the closer λ_2 is

to zero, the closer the graph is to being disconnected, which inhibits synchronisation in general.

5.5.3 Community effects in the burglary model

The work reviewed above, and the reasoning presented therein, reveals that the eigenspace of the Laplacian provides a link between the community structure of networks and the behaviour of a general class of dynamical processes taking place on them. In Section 5.4.2, the burglary model was shown to be of this type, and its behaviour was characterised in terms of the Laplacian eigenvalues and eigenvectors. A natural question, therefore, concerns how such effects are manifested in the burglary model and how they can be interpreted in real-world terms.

The concept of community which lies at the heart of such relationships has a relatively natural interpretation in the context of street networks. Communities, in the simplest terms, are groups which are well-connected internally but with comparatively few connections between them, and this definition maps almost immediately to the notion of ‘insularity’ of sections of the street network. Such sections would be expected to be characterised by some degree of redundancy, possessing a plurality of routes and a high level of within-group accessibility. The relative absence of connections between communities, on the other hand, would correspond to the idea that travel between such sections can only occur via a relatively small group of connections. While this characterisation is informal, it outlines the general principle of the community in terms of a spatial network.

Such notions of insularity can, of course, be reconciled with aspects of relevant criminological theory. Previous work which has considered the role of community structure in offending (Brantingham & Brantingham, 1975; Sampson & Groves, 1989) has typically identified units on the basis of demographic factors, but it is not unreasonable to hypothesise that these coincide, to some extent, with the physical structure itself. In addition, the community structure of a street network is deter-

mined primarily by planning decisions, and can therefore potentially be manipulated in order to realise policy objectives.

The question of which aspects of model behaviour should be analysed is one which requires some consideration. In the form presented in Section 5.4.4, the model simply evolves until a (known) equilibrium state is reached; the most natural interpretation of this is as a long-term average state, and so the path to equilibrium is of relatively little interest (or real-world relevance). Of greater significance is the behaviour at much shorter time-scales, and this can be explored by re-introducing a stochastic element to the model.

To re-cap, the model of burglary is governed by the equation

$$\frac{dQ_i}{dt} = \Gamma R_i - \omega Q_i + \eta \sum_{j \sim i} (Q_j - Q_i), \quad (5.61)$$

where $Q_i(t)$ is the dynamic component of burglary risk on segment e_i at time t , and $R_i(t) = S_i + Q_i(t)$ is the absolute risk which results when $Q_i(t)$ is combined with the static rate S_i . Substituting this gives the full expression

$$\frac{dQ_i}{dt} = \Gamma (S_i + Q_i) - \omega Q_i + \eta \sum_{j \sim i} (Q_j - Q_i). \quad (5.62)$$

In this formulation, the first term on the right hand side represents the boost effect of offending. Γ represents the size of the boost associated with a burglary incident, and $(S_i + Q_i(t))$ gives the rate at which offences take place. In order to convert this to a stochastic form, it is first written in terms of discrete time thus:

$$Q_i(t + \delta t) = Q_i(t) + \left(\Gamma (S_i + Q_i(t)) - \omega Q_i(t) + \eta \sum_{j \sim i} (Q_j(t) - Q_i(t)) \right) \delta t. \quad (5.63)$$

The boost term $\Gamma (S_i + Q_i(t)) \delta t$ is then simply replaced by $\Gamma \Theta(t)$, where $\Theta(t)$ is defined as the realisation of a Poisson-distributed random variable, with rate $(S_i + Q_i(t))$, in a time window of length δt . The discrete dynamics are therefore

given by the expression

$$Q_i(t + \delta t) = Q_i(t) + \left(-\omega Q_i(t) + \eta \sum_{j \sim i} (Q_j(t) - Q_i(t)) \right) \delta t + \Gamma \Theta(t). \quad (5.64)$$

In effect, the smooth boost process of the original model has been replaced by explicit instantaneous boosts, which occur probabilistically at each location according to the current level of risk given by the model.

In terms of the dynamics, the effect of the change is to split the process into a series of realisations of the original deterministic model, taking place between successive stochastic victimisations. Each time a burglary event occurs, the state of the system (taking into account the victimisation) represents an initial condition from which the original model runs unimpeded until the next victimisation, at which stage the cycle repeats. The behaviour of the model between victimisations is as would occur with a value of $\Gamma = 0$ in the original model, since the boost term is no longer present, and so the inter-event behaviour is governed by

$$Q_i(t + \delta t) = Q_i(t) + \left(-\omega Q_i(t) + \eta \sum_{j \sim i} (Q_j(t) - Q_i(t)) \right) \delta t, \quad (5.65)$$

with the initial condition for Q determined by the state at the time of victimisation. Whenever a victimisation occurs, the effect is to ‘reset’ the diffusion-decay model to reflect the state at the time of the new most recent incident. The location of the next victimisation is then determined by the subsequent behaviour of the model, since the rate of $\Theta(t)$ is a function of $Q_i(t)$. Because the next victimisation is more likely to occur on segments of higher $Q_i(t)$, it is by this mechanism that victimisation effects feed forward: enduring boost effects bias the rate of $\Theta(t)$, causing recently-victimised areas to be favoured as locations of the next victimisation. This is of particular interest since the diffusion-decay behaviour is expected to conform to the principles outlined in Section 5.5 concerning the relationship between dynamics and community structure. If this is the case, hierarchical synchronisation phenomena may cause identifiable community-level effects in terms of repeated victimisations

(or, in other words, hot-spots).

The mechanism by which such effects might arise can be understood in more concrete terms by considering the stage-by-stage nature of the diffusion-decay process, following the reasoning outlined in Section 5.5. It is straightforward to translate the equation (5.65) back to its equivalent continuous form

$$\frac{d\mathbf{Q}}{dt} + \omega\mathbf{Q} + \eta\mathbf{L}'\mathbf{Q} = 0 \quad (5.66)$$

which can be solved as shown in Section 5.4.2. Recalling that $(\lambda_h, \mathbf{w}_h)$ are the ordered eigenvalue-eigenvector pairs for the Laplacian, the solution is given by

$$\mathbf{Q}(t) = \sum_h q_h(0) e^{-t(\omega + \eta\lambda_h)} \mathbf{w}_h. \quad (5.67)$$

where the $q_h(0)$ are the coefficients of the eigenvector decomposition of $\mathbf{Q}(0)$. As shown above, the behaviour of (5.67) can effectively be broken down into stages as each eigenvector mode decays to zero. Defining τ_h as the ‘lifetime’ of mode \mathbf{w}_h (the time such that $q_h(0)e^{-t(\omega + \eta\lambda_h)} < \epsilon$ for a small choice of ϵ), distinct stages $[0, \tau_n)$, $[\tau_n, \tau_{n-1})$, \dots , $[\tau_3, \tau_2)$, $[\tau_2, \infty)$ can be identified (for the present purpose, $t = 0$ is taken to be the time of the most recent victimisation).

In the context of the stochastic model, all of these stages occur after every new victimisation. For the issue of hot-spot formation, the key question is during which stage the next victimisation takes place, since this defines the extent to which certain areas will be preferentially targeted. This can be illustrated by referring again to the diffusion system shown in Figure 5.11, which is similar in principle, and interpreting the line $t = 0$ as the time of the most recent victimisation. If, for example, the next victimisation takes place in the $[\tau_3, \tau_2)$ stage, the only variation in the risk is in the direction of the vector \mathbf{w}_2 (which corresponds to a division of the network into two components). The occurrence of the next victimisation is thus biased towards one of these components, so that there is a tendency for the same ‘half’ of the network

to be re-victimised. More generally, each stage $[\tau_{i+1}, \tau_i)$ is known to correspond to a situation in which the dynamics are synchronised, in some sense, within i communities, and so each of these communities would be expected to be unified from the perspective of re-victimisation.

Given these observations, a tendency for re-victimisation to occur within a certain stage implies, in the long-run, that the crime rate will be correlated at the corresponding community level. The tendency to fall within a certain stage, though, is determined by the underlying frequency of incidents; that is, the values of $S_i = (D + C_i)p_i$ and Q_i , which govern the rate of the random variable Θ . A higher static rate, for example, implies a shorter inter-event time and therefore a tendency for the next incident to occur during an earlier stage. In the context of the previous reasoning, this can be extended to a more general hypothesis: the long-term behaviour model involves synchronisation of crime activity within network communities, and the scale of these communities is determined by the underlying rate.

5.5.4 Numerical results

To investigate these phenomena, numerical simulations, based on equation (5.64), were carried out using a hypothetical network with stylised community structure. The network, shown in Figure 5.12a, was constructed so as to have a two-level community structure - a clear partition into two groups, each of which themselves split in two - and these communities are indicated in the figure. The network can be regarded as an augmented version of that studied previously (see Figure 5.9), where extra links have been added in order to make more explicit the community structure, as well as simply to enlarge the network in order to diminish the effect of stochastic variation.

Figure 5.12b shows the eigenvalues of the network in question. Most notably, clear gaps can be observed between the 2nd and 3rd eigenvectors, and between the 4th and 5th; this accords with the observations made in Section 5.5 concerning the re-

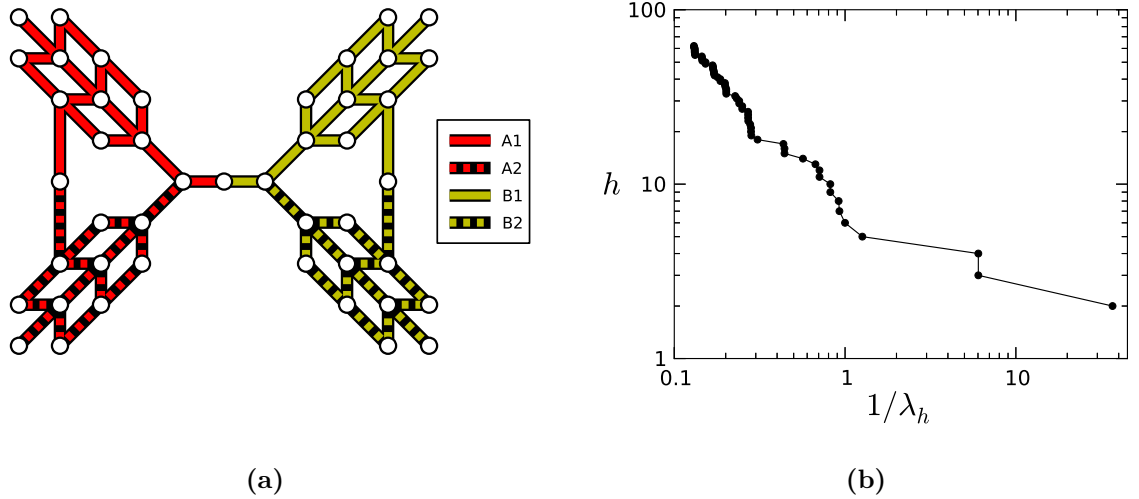


Figure 5.12: The hypothetical network used in analysis of community-level synchronisation. In a), the network is shown with links coloured according to community structure: red and yellow represent the coarse division into 2 groups, A and B , each of which can then be divided further (as shown by striped and plain links) to give a total of 4 communities. In b), the reciprocals of all non-zero eigenvalues of the Laplacian of the line graph are shown.

relationship between community structure and the eigenspectrum. It is these gaps which ultimately are expected to give rise to the dynamic behaviour of interest.

The purpose of the analysis is to consider the similarity in risk between network links over long-run simulations of the stochastic system. For any pair of links e_i and e_j , this is measured as the absolute difference in their dynamic risk values Q_i and Q_j , averaged over all discrete time-steps in the simulations. For a simulation which runs for a total of h time-steps, therefore, this similarity measure u_{ij} is defined to be

$$u_{ij} = \frac{1}{h} \sum_{k=0}^h |Q_j(k \times \delta t) - Q_i(k \times \delta t)| \quad (5.68)$$

Since the primary interest is in the relative sizes of these values, rather than their magnitude, the normalisation $\frac{1}{h}$ is not crucial, and simply allows simulations of differing length to be compared .

Computing u_{ij} for every combination of i and j gives a (symmetric) matrix of similarity values, \mathbf{U} , with dimension $m \times m$. This can then be visualised by plotting

a colour array in which the colour at each position represents the corresponding matrix value. Since the visual effect depends on the ordering of the links, the links are re-labelled prior to the computation of \mathbf{U} in order to group together those in the same community; in this way, any community effects should be manifested as blocks of similar colour.

For the numerical simulations themselves, several quantities can be varied: the parameters η , Γ and ω ; the initial condition for Q ; and the attributes of the network (*i.e.* the static risk, S). For clarity, and in order to isolate the diffusion phenomena, the static risk is taken to be homogeneous across the network, and its value $S_i = (D + C_i) p_i$ is set to be equal to a single quantity, S , for all i .

Simulations were run for a range of combinations of η , Γ , ω and S . In each case, simulations were run for 1,000 time units, using a discrete time-step δt equal to 0.01. The initial condition of Q was set to be uniformly equal to $\frac{\Gamma S}{\omega - \Gamma}$, its equilibrium value in the equivalent deterministic case, across all links.

The primary issue of interest is the variation in behaviour with S , with all other parameters kept constant: as discussed above, S determines the underlying rate of offending, and therefore the inter-event time.

Figure 5.13 shows similarity arrays for $S = 0.0004$, 0.004 and 0.08, in which the other parameter values are $\eta = 2.5$, $\Gamma = 0.005$ and $\omega = 0.01$. The three cases demonstrate how the community-level synchronisation is revealed at different scales as the underlying rate of victimisation is varied. In Figure 5.13a, two coherent blocks of colour are evident in the similarity array, corresponding to the two halves of the network, A and B . When victimisations occur more regularly, as shown in Figure 5.13b, the association between sub-communities (*e.g.* between $A1$ and $A2$) is much weaker, and the main association is within the 4 lower-level communities only. Finally, an extreme case is shown in Figure 5.13c, where the rate is sufficiently

high that community effects have almost broken down, and no well-defined blocks are present.

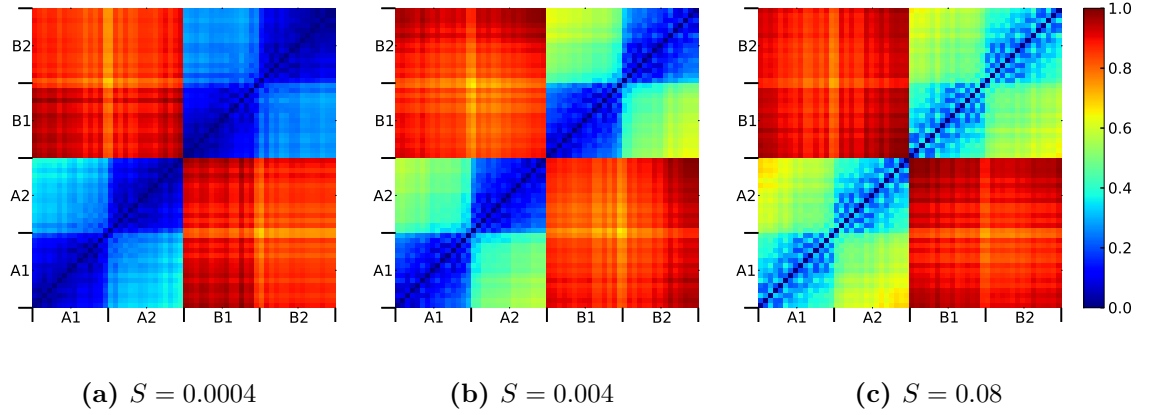


Figure 5.13: Inter-link similarity values for long-run simulations of the stochastic burglary model on the network shown in Figure 5.12a. The colour at each position represents the similarity u_{ij} between the pair of links e_i and e_j . The links have been re-ordered in order to group communities, and the values normalised in each case in order to facilitate comparison.

The community effects can be seen more explicitly by applying a threshold for the similarity values. Each pair of links is categorised according to whether their similarity is less than some value ν . Performing this for all pairs allows \mathbf{U} to be mapped to an equivalent binary array \mathbf{U}' , where

$$u'_{ij} = \begin{cases} 1 & \text{if } u_{ij} < \nu \\ 0 & \text{if } u_{ij} \geq \nu. \end{cases} \quad (5.69)$$

Figure 5.14 shows these arrays for the same parameter values as used in 5.13. Informally, the rank of these matrices (the number of independent/distinct rows) can be regarded as a measure of the number of independent synchronised communities. This can be seen in the matrices of Figure 5.14: a) shows small departure from rank 2, b) differs from rank 4 in a small number of cells, and c) is much larger.

Of course, when considered in the context of burglary, the systems shown differ in more fundamental ways than in their community-level behaviour. Changes in

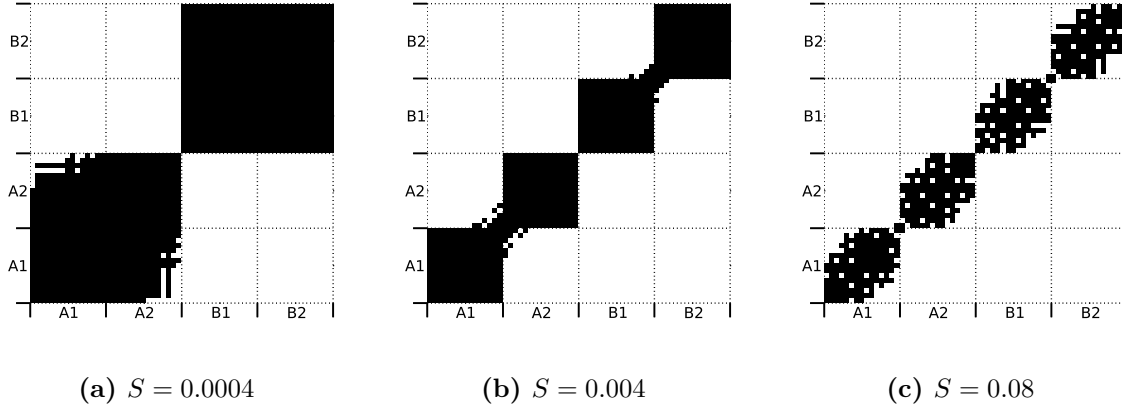


Figure 5.14: Binary versions of the similarity arrays presented in Figure 5.13. Each cell represents the value u'_{ij} , calculated using a threshold $\nu = 0.35$.

the static rate, for example, mean that the volume of offending varies considerably between the cases, and so the situations being represented are quite different. What the results do demonstrate, however, is the interplay between dynamical behaviour and community structure, and how the morphology of coherent hot-spots depends on the properties of the underlying social system.

This relationship can, however, also be viewed from the opposite perspective. In the simulations above, the network used was the same in each case and the static rate acted as the independent variable; this produced interesting behaviour by varying the correspondence between the static rate and network eigenvalues. In terms of this correspondence, however, an equivalent effect could also have been achieved by varying the eigenvalues; that is, by changing the network structure. A simple example of this, which further illustrates the connection between community stability and eigenvalues, can be seen by making a small modification to the network.

Figure 5.15a shows the network of Figure 5.12 with two links (and one vertex) removed. The links were chosen for removal because of their importance from the perspective of community structure: they provided the key ‘bridge’ between communities $B1$ and $B2$ and thereby contributed to the clarity of the 2-way partition. When the links are removed, the 2-level hierarchical community structure is less well-defined. While the 2-community left-right partition is still evident, a natural

3-way division can also now be seen, as highlighted in Figure 5.15a: groups $B1$ and $B2$ split, while A remains unified as a result of the remaining ‘bridge’ links. That this is a stable partition is evidenced by the eigenvalue distribution, shown in Figure 5.15b, in which the gap between the 3rd and 4th eigenvalues has become much larger. Following the previous argument, such gaps correspond to both the ‘quality’ of the partition, in some sense, and the dynamical behaviour on the network.

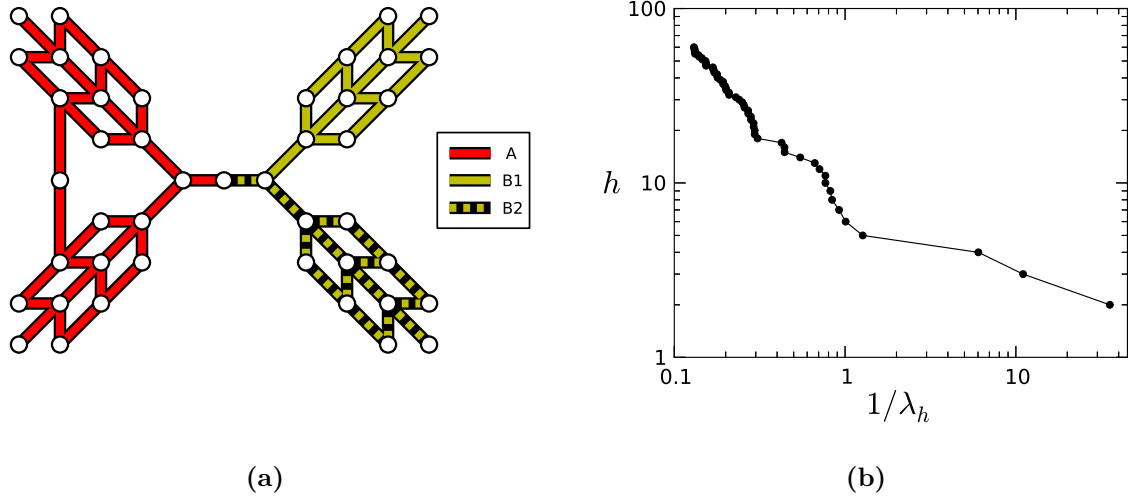


Figure 5.15: Modified version of the hypothetical network shown in Figure 5.12, after the removal of two links and one vertex. The natural 3-community partition is shown in a), and the extended gap between the 3rd and 4th eigenvalues can be observed in b).

Simulations carried out for this network reveal that the change in community structure does indeed have a clear effect on the similarity values observed. In Figure 5.16, similarity values are shown for simulations performed using the same parameter values as presented in Figure 5.13. It is clear that, in each case, the removal of the link has weakened the association between groups $B1$ and $B2$, so that the similarity between $B1$ and $B2$ is now noticeably lower in all cases than that within A . In all cases, the 3 blocks which correspond to the 3-way A - $B1$ - $B2$ partition are evident.

Such an association with community structure is as expected and, from a technical perspective, its emergence can be reconciled with the change in eigenvalues brought about by the link deletion. The fact that a much larger gap exists in the spectrum

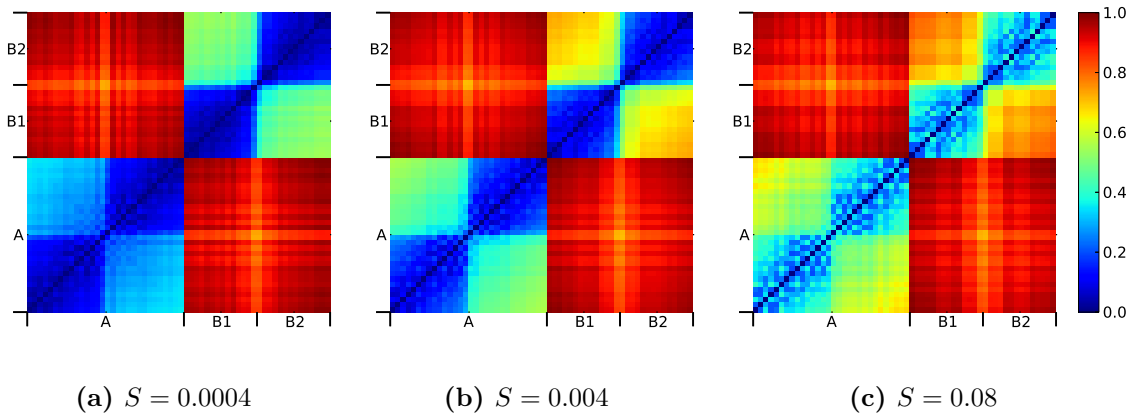


Figure 5.16: Similarity arrays for the modified example network, in which one vertex and two links have been removed. As before, the colour at each position represents the similarity u_{ij} between the pair of links e_i and e_j .

between the 3rd and 4th values means that the community structure identified by the 2nd and 3rd eigenvectors is reflected substantially in the dynamical behaviour. With respect to the wider point, this underlines the fact that structural changes to the network are capable of bringing about material changes in the shapes of hot-spots, even when the assumed criminal behaviour remains unchanged.

Although the results of Figure 5.16 demonstrate the addition of a coherent 3-way community structure, the effects of the 2- and 4-way partitions remain and can still be seen. Indeed, in general, the community relationships found in these simulations are not sharply-defined: they depend, to a certain extent, on threshold choices, and community membership cannot be specified definitively in some cases. This is a result of the stochastic nature of the dynamical system being simulated.

The stochastic behaviour causes these departures from ‘ideal’ community-level behaviour in two main ways, the first of which relates to the extent of variation between links. The occurrence of victimisation is, at any time, biased towards those areas which have higher attractiveness (which, by the eigenvector argument, will correspond to communities). However, since this is simply a bias, and victimisation is probabilistic, a non-negligible proportion of incidents will occur which do not follow the heterogeneity of risk exactly; lower-risk streets will still be victimised. For this

reason, patterns will only ever be approximate, and will only be revealed in long-run simulations (especially when the variation in risk is relatively small).

The second sense concerns the uncertainty in inter-event separation. The argument for community-scale reinforcement is grounded in the fact that, if a victimisation occurs within a given time window (*i.e.* stage) following a previous event, the distribution of risk can be characterised easily in terms of communities, because the variation is determined by only a few known eigenvectors. Of course, when victimisation is stochastic, there is no guarantee that events will fall consistently in such a time window; if not, the locational bias for the event will be different. Again, these effects only arise in the long-term average, when the dominance of a particular time-scale eventually becomes evident. With respect to analysis, such complexity essentially precludes a formal analytical treatment of the system, since the unpredictability in both these respects makes the understanding of time-scales, for example, analytically intractable.

To summarise, it has been shown in this final section of analysis that significant synchronisation can be observed in the behaviour of the burglary model, even when the dynamics are initialised stochastically on the basis of previous realisations. This not only extends known principles of network dynamics, but has real-world relevance in terms of the social system described by the model. The technical aspect and real-world implications will both be discussed in the final section.

5.6 Discussion

The aim of this chapter was to propose a mathematical model for the crime of burglary which was situated on a street network, and in which other aspects of urban form could be incorporated. The motivation for this was provided by the empirical results of Chapters 3 and 4, which suggested the existence of relationships between street network structure and both static and dynamic criminal phenomena. In particular, the findings that street segment centrality was positively associated

with burglary risk, and that post-victimisation risk diffusion could be understood in terms of network distance, suggested that the inclusion of street network effects was essential in order to capture a number of significant effects.

This aim could have been achieved by simply re-casting one of a number of existing models in a network framework; however, consideration of these models suggested that such an approach may not be optimal. In particular, the way in which offender behaviour - the primary mechanism by which the influence of the network is exerted - is conceptualised in those models does not take into account a number of aspects of theory; specifically, those related to purposeful or routine movement. Instead, therefore, a new model was proposed, in which network properties were used to reflect the criminogenic properties of, and relationships between, street segments.

After specifying the model, its properties were first investigated via numerical simulation. This demonstrated some of the basic properties of the model, and anticipated a number of potential practical uses: as a predictive tool, for example, or as a means of forecasting the effect of planning interventions. In addition, the effect of introducing spatial heterogeneity in burglary risk was shown, particularly in terms of its manifestation in long-term equilibrium states of the model.

Traditional mathematical analysis of the model was largely precluded by the complex topology on which the model was situated; since the complexity of the topology was one of the foundations of the model, this was unavoidable. Progress could be made, however, in relating precisely this topology to model outcomes. It was shown that a closed form solution for the model could be found in terms of the spectral properties of the network, thereby making explicit the link between network structure and model behaviour. Furthermore, this analysis clarified the way in which heterogeneities in static risk were manifested in the solution to the model, and identified a link with community structure in the network.

This community structure was the focus of the final part of the work, which concerned the behaviour of a stochastic version of the model and its relationship with community structure. This demonstrated not only that the behaviour of the model was synchronised, to an extent, within communities, but that the granularity of the community structure in question was determined by the underlying stochastic rate of victimisation. This is a finding with implications for urban design, which will be explored below.

5.6.1 Modelling outcomes

The model of burglary was constructed with two primary objectives in mind: to situate it on a network, and to incorporate relevant theories of offender behaviour. Without the latter, the value of the model would have been somewhat undermined, since no existing hypothesis suggests a relationship between crime and networks *per se*; all findings rest on their role in shaping the activities of offenders. Particular examples of these, such as routine activities theory and pattern theory, emphasise the role of purposeful travel through urban areas in shaping awareness spaces; however, the inclusion of such principles in models of crime has, thus far, been limited to agent-based approaches.

In the model proposed here, it was suggested that it is not, in fact, necessary to model individual behaviour explicitly (*i.e.* using agents) in order to incorporate such concepts in a model. Indeed, the primary innovation here was to note that these principles can be captured in a traditional mathematical model by making use of the non-trivial topological properties of the spatial setting. The work of Chapters 3 and 4 showed that burglary phenomena can be reconciled with network metrics: the inherent risk of street segments is predicted by their regularity of use (measured via betweenness), and the diffusion of risk corresponds to network proximity.

Because of these observations, the model which was built did not include an explicit description of offender movement; instead these concerns were encoded via

the structure and properties of the underlying network. In short, the reason for this is that the properties of space which are of relevance to crime do not change over time: the inherent level of risk is known *a priori*, as are the directions in which additional risk is likely to diffuse. Modelling offenders directly therefore adds little, and simply has the effect of introducing unnecessary complexity. In fact, their absence corresponds more closely to the kind of situation which might be encountered in practice: when an offence occurs, the identity of the offender is typically not known, so that the question of a potential follow-up concerns a generic, rather than specific, offender.

In terms of modelling more generally, this demonstrates that the inclusion of additional granularity, in a modelling sense, is not always preferable; in the case of crime, all issues related to offender movement are captured by the properties of the spatial setting. This does, of course, rely on the spatial structure being sufficiently complex to give rise to interesting features, and this strengthens the case for the use of realistic environments in models of crime (and social systems more generally). In this sense, a parallel can be drawn with epidemiological models. The assumption in those models of a ‘fully mixed’ population was shown to be flawed when it was shown that results can differ materially when social/contact networks are taken into account (Moore & Newman, 2000; Pastor-Satorras & Vespignani, 2001); in the case of urban spatial systems, street networks are a similarly significant structure.

5.6.2 Mathematical outcomes

The general principles encoded by the mathematical model - a combination of self-excitation and simple diffusion - are not novel, and have been inspired by similar previous models; instead, the primary innovation concerns their situation on a non-trivial network. Accordingly, the technical novelty of the chapter concerns this issue: the relationship between dynamical properties and network structure. In this respect, the key finding is the link between model dynamics and the spectra of the network on which they take place; this, in the first instance, allows a closed form to

be found for the dynamics in terms of the eigenvalues and eigenvectors of a related graph object.

This link, coupled with the fact that graph spectra are known to correspond to their community structure, allowed the equilibria of the model to be characterised in these terms. Among other things, this demonstrated the extent to which heterogeneities in static risk were ‘smoothed’ in the equilibrium: sharp discontinuities were moderated to a greater extent than those which were well-aligned with strong communities.

The latter element of the chapter built on this, by examining the behaviour of the model under stochastic input, and this represents the primary technical novelty of the analysis. Although the relationship between dynamical behaviour and community structure has been studied elsewhere, it has typically been concerned only with synchronisation-like behaviour of deterministic systems, running from known initial conditions with no additional input. As outlined in the final section, stochasticity - and, moreover, its role in the regular re-setting of the system - presents significant additional challenges. In fact, the fundamental question is transposed to one of emergence, since the overall system here is a series of deterministic realisations, each initialised probabilistically according to the evolution of the preceding one.

The results here demonstrate that community behaviour can still be seen in such circumstances; that is, that the relationship is sufficiently strong that the expected associations can still be seen in a stochastic scenario. In addition, the effect of changing time-scales (defined here by the underlying static rate of victimisation) has been shown to act as a tuning parameter between scales of community structure, acting by varying the stage of evolution at which the diffusion behaviour is re-initialised. This is not an aspect of network dynamics which has been investigated previously, but represents both a realistic case and one which can be reconciled clearly with theory.

5.6.3 Real-world implications

The most natural practical application of the model presented in this chapter is its potential use as a predictive tool; that is, as a means of producing short-term forecasts of the spatial distribution of burglary risk. The value of predictive modelling in general has been outlined elsewhere and requires little further comment; again, however, the network setting is the aspect of this contribution which sets it apart. There are two primary senses in which this can be expected to improve the utility of models. The case for the first of these is made by the preceding two chapters: the model incorporates effects which have been shown to be statistically significant, and therefore ought to represent an improvement upon existing methods, in terms of predictive capability. The second sense relates to practical deployment, and concerns the fact that interventions, such as patrolling, would typically be specified in terms of street segments. Predictions specified in those terms are therefore of immediate use, and could eventually be used in the efficient design of entire patrol routes.

The other property of the model which was investigated - the observation of unified behaviour within network communities - can also be translated into real-world terms. To do so requires two key links to be made with the real-world situation: the meaning of community structure in terms of urban design, and the meaning of the behavioural variations explored in the simulations. The first of these is relatively simple. In line with the earlier argument, community structure in street networks relates to the partitioning of the network in a manner such that there are relatively few connections between sections, but relatively many within. This relates in a straightforward way to an intuitive notion of ‘insularity’ in neighbourhoods, and features such as housing estates would be expected to be identifiable on this basis. Importantly, though, this definition does not apply to only one scale, since community structure can typically be identified at more than one resolution: to give an extreme example, the city of London could be split conveniently into two parts by the Thames river, but a more granular division could also be envisaged based on its

many metropolitan population centres. In terms of urban policing, it may be the case that neighbourhoods, or beats, show a degree of correspondence with community structure. As shown in Section 5.5.4, the coherence of each scale is determined partly by the underlying rate of victimisation.

The implication of the numerical results of Section 5.5.4 is that a much greater association in the risk of crime exists within communities than between them. More specifically, burglary risk remains relatively uniform within such areas, so that they can be taken to be equivalent, to some extent, in terms of burglary risk; instead of specifying risk in terms of individual streets, communities of streets can be identified and regarded as single entities. It should also be emphasised that the risk in question is dynamic risk; that is, risk which is representative of repeat or near-repeat victimisation effects. From that perspective, the implication is that follow-up incidents are likely to occur within the same community as the previous incidents which led to the elevation in risk. This suggests that community structure should be taken into account when determining the spatial extent of interventions intended to prevent near-repeat incidents.

5.6.4 Further work

A number of avenues for further work can be identified, both in terms of model refinement and practical deployment. One possibility which is immediately apparent is the use of more sophisticated measures of static risk. Only betweenness is considered here, but the possibility remains of incorporating heterogeneity in the inherent attractiveness of properties on particular streets, or, indeed, the use of more nuanced versions of betweenness. Refinements in this sense have the potential to better reflect the true nature of static risk.

Aside from measures of static risk, the nature of diffusion could also be changed. In Chapter 4, it was shown that the diffusion of burglary risk is not uniform, and that its spread can be predicted by a pair-wise measurement defined as commonality.

This principle could be incorporated relatively easily: instead of risk diffusing at an equal rate to all adjacent street segments, the rate for each segment could be taken to be proportional to commonality. This was not done here because of the complex nature of commonality - its values cannot be characterised simply, and so mathematical analysis would have been impossible - but could be included in future numerical implementations.

One other aspect which is absent from the model is the effect of policing: its effect on crime, and the behaviour of police officers. This could be incorporated in a number of ways, including the direct suppression of burglary risk and a tendency to allocate resources to hot-spots, but significantly more empirical research is required before this can be done in a well-grounded way. At present, neither police behaviour nor its effect on crime are well-understood, and so any modelling assumptions would be difficult to justify.

A final point concerns practical implementation. If the model is to be used in practice, either as a predictive tool or for the testing of policies, some form of calibration to real data will be required. This will entail a number of technical tasks: the identification of suitable measurements of static risk; the resolution of real-world crime patterns into static and dynamic effects; and the selection of appropriate parameter values. Each of these represents a significant challenge, but all are necessary if the model is to be translated from its hypothetical setting to a real-world implementation, and it is this area which represents the primary avenue for further development.

Chapter 6

A model of the London riots

6.1 Introduction

The modelling work of the previous chapter, and the majority of the analysis in the thesis to this point, has been concerned with the crime of burglary. As noted previously, burglary is a convenient example for such study, for both practical and theoretical reasons. The latter of these is primarily due to the fact that the relevant criminological theory is relatively well-developed and incorporates several principles which can be expressed mathematically (or, at least, correspond to mathematical concepts). These general theoretical foundations are, however, common to a number of crimes, and so it might be expected that those could also be modelled using the same general approach, with details modified. This is the case for crimes which share a number of characteristics with burglary: high-volume offences, occurring in urban areas, typically committed by a single actor without a target-specific motive.

This definition certainly does not, however, encompass all crime, and the differing characteristics of other contexts require different modelling approaches to be taken. The purpose of this chapter is to demonstrate this by considering a crime type which contrasts significantly with burglary: large-scale civil disorder, or rioting. This is made possible by the availability of data for one such example, the London riots of August 2011, and that episode is used as a case study here. Riots are sufficiently rare events - and each can be regarded as unique in a number of ways - that to model them in generality would be infeasible, and the choice to consider

one particular episode is made for that reason. Nevertheless, the possibility remains that general principles established for the London case could apply more generally.

Several aspects of rioting render it distinct from other crime types which have been modelled previously. The simple fact that riots are rare events, each of which is subject to particular contextual peculiarities, means that there is no coherent general theory concerning those aspects which may be the subject of modelling, such as target choice. For example, prior to recent work using data from the same London disorder examined here (Baudains *et al.*, 2013b), few studies had addressed the question of whether the patterns observed in such exceptional circumstances were consistent with more general theories of criminology, such as pattern theory (Brantingham & Brantingham, 1993a). Although this appears broadly to be the case, significant differences are still present in terms of the particular mechanisms at work.

While it may be argued that much of the offending associated with riots is opportunistic in nature (*e.g.* looting), this is in a very different sense to how burglary, for example, can be regarded as such. Whereas opportunities for burglary are, it is argued, typically encountered and taken during routine activities, opportunities for riot-related disorder occur precisely because of the extreme (*i.e.* non-routine) circumstances which have arisen. For this reason, relatively specific theories concerning, for example, the influence of urban form on offending patterns, cannot necessarily be assumed to apply in such cases. Instead, it may be that such influences are simply overridden during a riot; certainly, the decision-making process is highly dependent on the precise situation and on the activities of others.

This points to the other main factor which sets riot apart from other crimes: the fact that such crimes are fundamentally a collective action. In many ways, offending is made possible when normal forms of guardianship are overwhelmed by the volume of participants; because of this, the attractiveness of a target is linked intrinsically to

the number of people seeking to offend against it. Participants influence each other and, because of the ‘safety in numbers’ principle, there is a tendency for crowds to unify in their collective target choice. Together, these factors mean that riots represent a very different modelling proposition to more common, habitual crimes.

The practical importance of understanding such events is, however, substantial. Riots cause considerable damage, in financial terms and otherwise: while this may be less, in the aggregate, than more common crimes, the fact that it is rare and unanticipated, and that the form of damage is non-standard (and serious, in terms of personal injury), means that the effects are magnified. Such events also attract extensive publicity and associated scrutiny, and can become significant politically.

Many of the policy questions which arise in the context of riots concern their policing, and how such an incident can best be managed and suppressed. The rarity of such incidents means that this is particularly challenging, for a number of reasons. Firstly, it implies a paucity of data concerning previous incidents; moreover, the idiosyncrasies associated with each episode mean that it is far from certain that findings are generalisable. In addition, there is little scope for the police to plan for such events in an evidence-based way, or to do so based on experience, since ‘live’ scenarios are so rare and demand exclusive focus on rapid suppression.

It is in this respect that the modelling work described here has the potential to contribute. To answer questions related to policing during such an event requires understanding of the spatio-temporal patterns of riots and how they develop over the course of an episode. Models which relate to this aspect have been shown to have potential value for other crimes (see discussion in Section 5.2), and a similar approach is explored in this chapter. Because of its focus on the development of an ongoing riot, this work is distinguished from the majority of other riot related research: rather than exploring the causes of an incident, its initiation is taken as the starting point and the focus is on the subsequent evolution. The particular

value of modelling in this respect is in its ability to explore policy questions in a quantitative way; not only might results suggest a policy change, they can indicate *how much* the policy ought to be changed. In matters of resourcing, for example, such recommendations can play a key part in the evidence base.

In order to inform the model, a number of empirical analyses are first carried out using data from the London disorder. Although this is purposely done in relatively coarse detail, a number of inferences concerning the behaviour of offenders can nevertheless be made. These relate to different aspects of the disorder: the time-course and development of events, spatial patterns in participation, and the choice processes of offenders.

Motivated by these ‘stylised facts’, the model itself is composed of three parts, notionally corresponding to the stages of an offender’s participation. Each of the stages is modelled using analogies with other fields, such as epidemiology and retail, so that the final system is a hybrid of several pre-existing models. This is then investigated numerically in order to examine its qualitative agreement with observed data. Finally, with its potential use as a policy tool in mind, the model is used to examine particular questions of the type that have been raised in discourse following the London riots.

6.1.1 The London riots

The London riots occurred between the 6th and 10th of August 2011, and constitute the most widespread and sustained period of civil unrest in the city for at least 20 years. The apparent catalyst for the disorder was the fatal shooting by police of a suspect in Tottenham, North London, and a subsequent protest; however, much of the offending that followed was not overtly linked to this. The subsequent disorder spread to a number of locations across the extent of London, and included a substantial volume of acquisitive offending alongside more politically-motivated unrest.

Overall, the events were characterised by repeated episodes of looting, rioting, arson and inter-personal violence, all of which were widely publicised in the media. In terms of their consequences, numerous instances of injury were reported, including five deaths, and extensive property damage was incurred, for which liability has been estimated as £250 million (Metropolitan Police, 2011). In addition to that occurring in London, similar, though significantly smaller-scale, episodes occurred in other UK cities, such as Manchester and Birmingham, seemingly inspired by events in London. For reasons of data availability, however, only the case of London is considered here.

In their aftermath, the London riots have been the subject of much research in the academic (*e.g.* Gross, 2011; Baudains *et al.*, 2013a), governmental (Riots Communities and Victims Panel, 2011) and journalistic (The Guardian and LSE, 2011) communities. In particular, a number of inquiries have examined the riots from particular perspectives, such as that of the police (HMIC, 2011) and that of the citizens of London (Riots Communities and Victims Panel, 2011), and addressed a number of questions of policy in their reports. Much of this research is concerned with the social conditions which created the environment in which the riots could take place, as well as the individual motives of rioters during the incidents themselves. These refer variously to factors such as unemployment, poor police relations and gang membership, though more fundamental issues relating to inequality and trust of government are also raised.

Beyond these questions, one prominent theme of the subsequent analysis of the riots concerns the response of the authorities. Although order was restored after five days, the possibility remains that the riots could have been brought to a swifter conclusion, had the police acted differently. Indeed, it is open to question whether the suppression can, in fact, be attributed to police activity; other possibilities, such as that the riots had run their natural course, or even that a change of weather was responsible, have also been put forward (Riots Communities and Victims Panel,

2011). It is clear, however, that the activities of the police did materially influence the evolution of the disorder.

As regards particular elements of police response, criticism focussed on three aspects in particular: that the police were inadequately prepared, that they were slow to react, and that they employed inappropriate tactics during their interactions. Such shortcomings were acknowledged by official inquiries (Metropolitan Police, 2011, 2012; HMIC, 2011), and many of the identified areas of concern relate to these general themes. Alongside a need to anticipate better the disorder itself, the inquiries emphasise the need to establish a level of policing resource, and mode of response, commensurate with an outbreak of this magnitude. However, very few of the reports (at least those in the public domain) make reference to how such estimations might be made; rather, such assertions are made in general terms. Given the aforementioned rarity of episodes of this type, there is little scope for refining such estimates empirically. This underlines the value of obtaining quantitative recommendations by other means.

6.1.2 Previous research on riots

Rioting has been the subject of relatively little research within criminology, and the majority of recent modelling effort has also been focussed on other crimes. Nevertheless, the limited body of work which does exist provides partial justification for the modelling approach used in the remainder of the chapter, and so will be briefly outlined here.

The majority of criminological research concerning rioting has concerned the decision-making of offenders, particularly in relation to the collective dimension of offending. Early theorising by Freud (1957) and Le Bon (1960) suggested that riot participation was driven by irrational behaviour, with individuals engaging in “animal-like” activity. According to this, target choice is essentially random and the crowd as a whole can be regarded as being driven by an irrational collective mind.

This assumption of irrationality has, however, been contradicted by a significant volume of subsequent work. Several researchers (Berk, 1974; Mason, 1984; Auyero & Moran, 2007) have argued that participants do indeed act rationally, subject to the caveat that this may be limited to some extent. In particular, it is claimed that offenders evaluate (however briefly) the relative merits of possible courses of action - whether to offend, or particular targets - by taking into account such factors as the potential reward available and the distance to be travelled. In addition, one factor which is hypothesised to play a significant role in this decision during rioting is the likelihood of capture (McPhail, 1994).

Assuming that these actions are indeed rational, it is natural to hypothesise that more general criminological theories might also apply; these include crime pattern theory (Brantingham & Brantingham, 1993a) and social disorganisation theory (Shaw & McKay, 1969), which have been discussed previously in this thesis. Though these cannot be assumed to apply *a priori*, research by Baudains *et al.* (2013b) supports the hypothesis that such principles were evident in the London riots. In particular, the results of the conditional logit model (as used in Chapter 4) used suggest that distance was a significant impediment to offending, and that the presence of retail facilities was associated with greater offending.

As well as target choice, such principles might also be extended to apply to the initial decision to offend. Theory states that environmental precipitators (in this case, knowledge of ongoing rioting) can serve to prompt, pressure, permit or provoke offending (Wortley, 2001); the ‘safety in numbers’ effect, in terms of the risk of arrest (Granovetter, 1978; Epstein, 2002), is a particularly clear example of how this might influence cost/benefit calculations in a riot scenario (Myers, 2000). In the case of London, it has been claimed (Gross, 2011) that awareness of disorder provided a self-reinforcing stimulus to rioter involvement, facilitated in many cases by social media. These ideas can be expected to play a role in the evolution and

spreading of disorder, and suggest a contagion-like mechanism.

The limited number of attempts which have been made to model riots have employed both continuous and agent-based approaches. The former have, in general, been attempts to adapt models of crowd dynamics to the case of rioting, while doing little to accommodate realistic human behaviour (Kirkland & Maciejewski, 2003; Bhat & Maciejewski, 2006). Agent-based models, following the civil violence model of Epstein (2002), have had some success in using game-theoretic concepts to inform agent behaviour (Yiu *et al.*, 2003; Goh *et al.*, 2006). These have generally, though, given little attention to geographical concerns, either treating movement as random or else in a fairly naive sense. More recent approaches have remedied this somewhat by incorporating real spatial data via GIS (Weidmann & Salehyan, 2013) and including sophisticated spatial decision-making (Torrens & McDaniel, 2013). In general, though, the incorporation of such behaviour has been limited to agent-based work, and has not been grounded explicitly in empirical observation.

6.2 Empirical analysis

Although the analysis of previous riots has succeeded in providing insight into various aspects of the behaviour of rioters, relatively few studies of events prior to those in London have contained substantial quantitative analysis. This contrasts with the case of burglary, for example, for which criminological theories are supported by a large body of quantitative evidence. Although the aim of the current work is modest, in terms of the sophistication of the model - offender decision processes will necessarily be formulated in a very coarse way - it is, nevertheless, essential that it should be built, at least partially, on the basis of empirical observation.

As noted in the previous section, it has been argued that individuals act, to some extent, rationally during a riot (*e.g.* McPhail, 1991); that is, that they engage in some form of heuristic cost/benefit analysis in the course of their decision-making. Assuming that this is the case, a modelling approach which seeks to encode these

decisions, and the factors which contribute to them, is a viable one. The most natural way to understand such processes is by examining the revealed preferences of rioters, and therefore analysis of data from the London disorder represents an important preliminary step in the analysis. In the pursuit of parsimony, this is done in a relatively simple way, but more sophisticated analysis of the same data can be found elsewhere (Baudains *et al.*, 2013b).

6.2.1 Data

The analysis of offender behaviour is based on examination of crime data provided by the Metropolitan Police Service. This dataset contains the details of all individuals arrested in the period between the 6th and 11th of August 2011 for an offence classified as being associated with the riots. Each record relates to a single incident, and contains a number of details, including the area in which the offence took place and a point estimate of the date and time at which it occurred. In addition, many of the records also contain information about the offenders themselves, such as their ages and home addresses. The latter of these is crucial to the analysis: since it represents, in some sense, the origin of the offenders, it provides information about the entire process of participation (*i.e.* the cycle from start to finish).

The use of police crime records in analysis is typically regarded as being subject to a number of caveats, and these are particularly acute in the case of the riots. Among these is that crime recorded by police is perceived to be significantly less than the total volume of crime, either because the crime is not reported or because no suspect is arrested. In the case of London, such was the volume of offending and lack of order that a large volume of incidents are likely to either never have been recorded or assimilated into the general body of offending; certainly, many participants were never identified (Home Office, 2011).

This issue is exacerbated by the fact that the records may not be a representative sample of total offending. There may be systematic differences between offenders

who were arrested and those who were not, as is speculated to be the case for other crimes. Indeed, this may even vary spatially: if police were successful in gaining control in one location, it may be that offenders in that location are over-represented. These issues are unavoidable with this dataset; nevertheless, it is the best available and its analysis does offer some insight into the patterns present.

A number of features of the dataset are also worthy of note from a technical point of view. Firstly, no offender appears more than once in the dataset, so that each participant is associated with only one crime. This fails to reflect the fact that some offenders committed multiple crimes while participating; however, this fits the model well. The model is built around individuals' spells of participation, so that the primary issue is simply where these took place, rather than the volume or type of offences which took place. Indeed, counting with multiplicity might overstate the activity of individual offenders and bias the data.

The quality of the spatial data is also a cause for concern. Firstly, of the 3,914 records present, only 2,299 contain entries for both residential and offence location; in the remainder, at least one of these was not determined or recorded by police. Since both of these elements are required to calculate some quantities, such as the journey-to-crime, the analysis is based on the complete records only. The specificity of the spatial data is also variable; in some cases, the location is given in terms of postcode only, while others are much more precise. In order for this to be kept uniform, therefore, the points are first mapped to the centroids of basic census units, and distances calculated using those. Although this procedure introduces noise, it is the only method which can be applied consistently across the data.

The spatial unit which is used in the analysis (and following model) is the census Lower Super Output Area (LSOA), which is one level of a hierarchical geographical structure defined by the UK government (Office for National Statistics, 2011). LSOAs are used because they represent a compromise between smaller units, for

which data would be sparse and which may be unwieldy in the model, and larger ones, which would imply a loss of detail. In addition to the geometries themselves, a number of statistics are also considered at LSOA level. As well as the residential populations, taken from the census itself, an indicator of social deprivation is used to provide context. This is the Index of Multiple Deprivation (IMD), which was produced by the Department for Communities and Local Government (2010) and can be used in conjunction with the census. It is a UK-wide indicator, by which areas are ranked according to a combination of employment, health, education, housing, and other such factors.

The final set of data used concerns one type of location which is identified as being particularly relevant to the analysis: sites of retail activity in London. For this, a set of locations defined as ‘retail centres’ or ‘retail cores’ by the Department for Communities and Local Government (2011) is used. The sites identified within this are consistently-defined ‘areas of town centre activity’, and measurements of the total area of retail floorspace are given for each. There are 246 such sites for London as a whole, spread across the extent of the city.

6.2.2 Riot characteristics

The first general observation is one which is not, in fact, derived from the offender data, but is one of the few quantitative findings contained in the various subsequent government reports. It concerns the predominance of acquisitive crime within the disorder - primarily in the form of looting - and, in particular, the disproportionate targeting of retail premises. Crimes against commercial premises, including both acquisitive crime and criminal damage, accounted for 51% of all offences in the UK as a whole (Home Office, 2011): a significantly higher proportion than any other individual crime type. Since London accounts for the large majority of offending, this figure is likely to be valid for that particular case also.

Qualitative evidence within these reports also supports this view: a number of areas

of dense retail activity, such as Clapham Junction and the town centres of Croydon, Ealing and Brixton, are cited as the most intensely-targeted areas of London. In addition, reports of offender accounts refer to the prospect of looting as a key driver of participation. This can be reconciled with crime pattern theory (Brantingham & Brantingham, 1993a); the richness of opportunity at retail premises is likely to be common knowledge amongst riot participants, and they therefore act as *crime attractors*.

Journey to crime Moving on to consider the offender data, the first stage involves simple journey-to-crime analysis of the type frequently performed in criminology (*e.g.* Townsley & Sidebottom, 2010). This is done by comparing the home addresses of offenders with the locations of their offences, thereby examining the distances they travelled in order to offend. As mentioned in the previous section, calculations are performed using the centroids of the LSOAs in which the relevant locations lie, and a simple Euclidean distance is calculated. Although it is acknowledged that this concept of distance may not represent the true ‘cost’ of travel, as perceived by a rioter, it is the only meaningful metric which can be applied consistently to the data available. Furthermore, there is considerable doubt as to whether alternative approaches, such as those incorporating travel time, can be assumed to apply in such extraordinary circumstances (particularly the disruption of public transport).

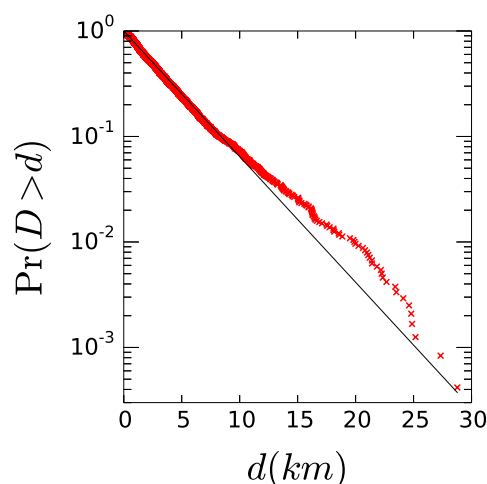


Figure 6.1: The ‘journey to riot’: the distribution of D , the distance between offenders’ home addresses and their offence locations, measured in kilometres, is shown as a complementary cumulative distribution.

The results of this analysis are shown in Figure 6.1, in which a clear distance-decay relationship can be seen. This is consistent with expectation, and with other studies of the London riots, which have found that offenders tended to target locations close to their homes. As well as visual inspection, statistical methods are available by which the observed distribution shown in Figure 6.1 can be tested for adherence to a number of decay distributions, such as power-laws. Although these tests suggest that the distribution cannot be said to have arisen from any of these, the best fit is found for an exponential distribution with parameter 0.274 (for which the Kolmogorov-Smirnov distance between the distributions is 0.0246). That the distributions do not conform exactly is unsurprising; offenders' perception of distance is determined by a complex combination of factors (including some which are temporally-varying). Noise is also introduced by the use of LSOA geography.

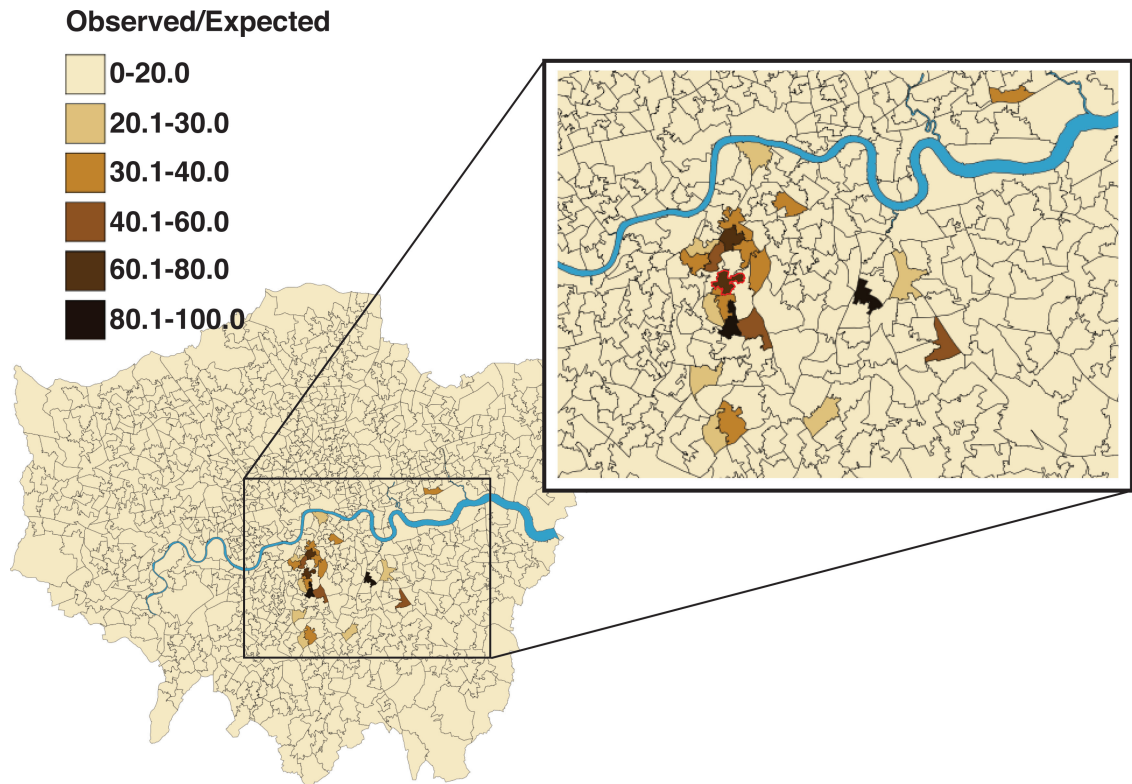


Figure 6.2: The home locations of rioters whose offences occurred in Brixton, shown by colouring LSOAs according to the proportion of the residential population who offended.

In order to present a concrete example of the distance decay principle, Figure 6.2 shows the origins of rioters whose offences occurred in Brixton. It is evident that almost all participants in the Brixton offending lived in close proximity to their targets, and gravitated towards the central retail area. Although the trend is not uniform, it is illustrative of the general pattern of distance decay evident in the data.

Relationship with deprivation The origins of offenders and the eventual offence locations can also be analysed independently. Of particular interest is the issue of whether there exists any spatial pattern in levels of participation; that is, whether some areas provided more rioters than others. In light of the distance decay relationship observed, this would have clear implications for the evolution of disorder, since targets close to areas of high participation would be at greater risk. Informed by the hypothesis, put forward in several reports (*e.g.* Home Office, 2011), that areas of higher deprivation were more susceptible to rioting, this issue is examined by comparison with a measure of deprivation, in the form of the IMD. Two of the factors incorporated in the IMD, child poverty and youth unemployment, have previously been identified by Ben-Galim & Gottfried (2011) as factors contributing to the riots.

Figure 6.3a shows the variation in levels of participation with the IMD of the LSOAs of London. Anticipating its use in the following model, a score is derived from the IMD (which is rank-based) by normalising the ranks to the range $[0, 1]$, so that 1 represents the most deprived, and this is plotted on the horizontal axis. The values plotted are the proportion of the population that offended, for equally-sized sets of LSOAs grouped according to their ranking. It is evident that, in general, higher deprivation is associated with increased offending: the relationship is almost uniform, and participation varies by approximately one order of magnitude across the full range. This suggests that IMD is indeed a valid predictor of the tendency to offend; this is convenient, since it summarises a number of demographic factors.

The hypothesis that offending itself was also biased toward more deprived areas is also supported by the data. Figure 6.3b shows the growth of the cumulative

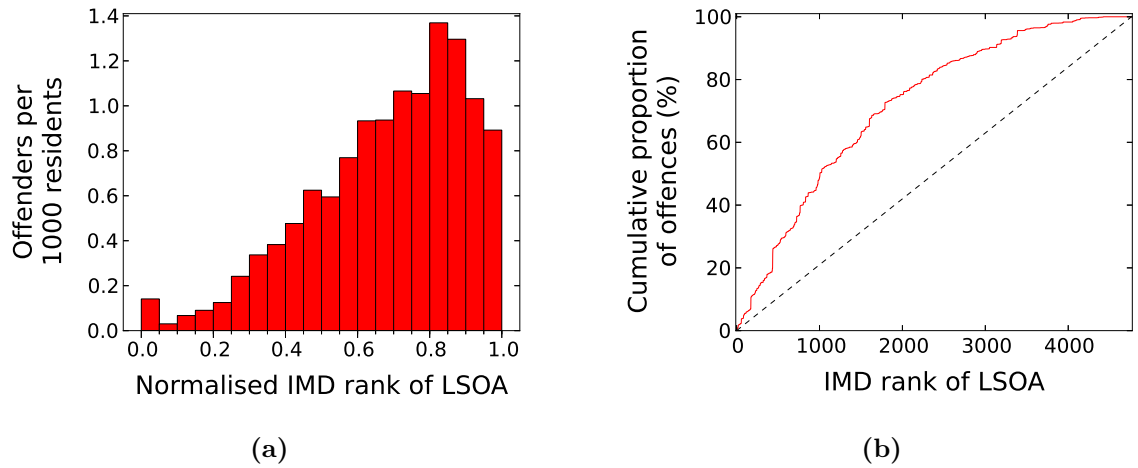


Figure 6.3: The relationship between offending and deprivation. Figure a) shows levels of participation amongst the residents of groups of LSOAs, ordered according to a score derived from their IMD rank. In b), the cumulative number of offences when LSOAs are ordered according to IMD is shown in red, while the dashed line represents a hypothetical uniform distribution. The gap between curves shows that offences occurred disproportionately in more deprived areas.

proportion of offence locations when LSOAs are ordered according to deprivation; in this case, the zero-ranked LSOA is the most deprived. It is clear that the total volume of offending is concentrated in the most deprived areas, with approximately 50% of the offences occurring within the 20% most deprived areas, for example. This supports the general conclusion that more deprived areas experienced greater disorder, both in terms of participation and offending. Considering this in conjunction with the journey-to-crime results does raise a question of causality, since the fact that deprived areas act disproportionately as both origins and destinations will influence the distance distribution, and *vice versa*. Although this is not explored here, more detailed analysis by Baudains *et al.* (2013b) suggests that both effects remain when controlling for the other.

Temporal patterns One final issue of note concerns the temporal patterns within the data, both within individual days and across the period as a whole. Figure 6.4 shows the distribution of offences through time, based on the recorded time of incidents. A number of qualitative features are immediately apparent: firstly, that the period as a whole can be broken down into a series of daily episodes, separated by periods of negligible offending. Within each particular day, the majority of events

are seen to occur at night, building to a general peak around midnight and the early hours before decreasing to very small levels. While various explanations for this have been put forward (Riots Communities and Victims Panel, 2011), a particularly compelling one suggests that awareness of disorder provided a self-reinforcing stimulus to rioter involvement (Gross, 2011), and that an element of social contagion was in effect.

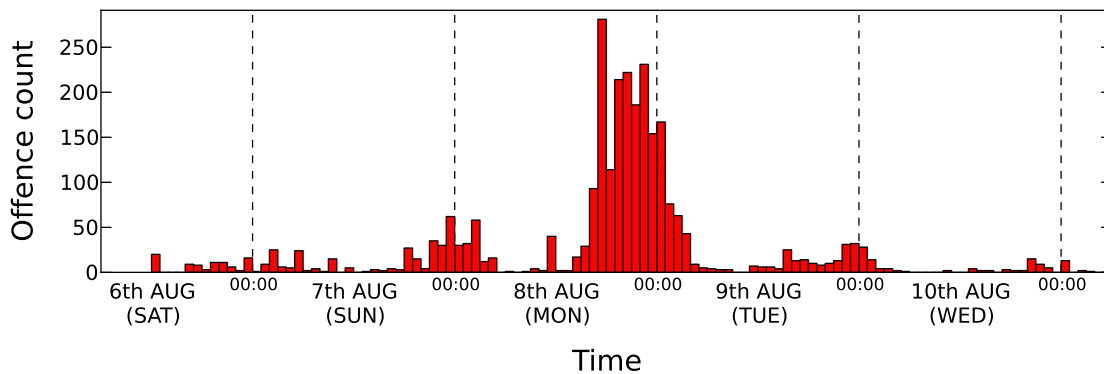


Figure 6.4: Time series of offending across the period of the London riots, with offences grouped by hour.

Noticeable changes can also be seen from day to day. After a modest increase in overall offending on the second night, a marked increase in overall offending is seen on the third night, before an equivalent decrease on the fourth night. This escalation was reflected, though not mirrored, by the volume of police resources operational on each day (Metropolitan Police, 2011), as shown in Figure 6.5.

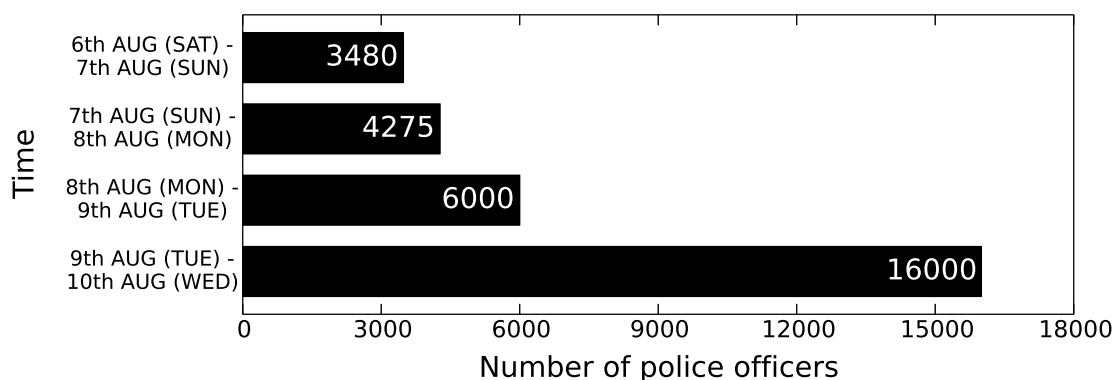


Figure 6.5: The numbers of police officers deployed in London on each of the four main nights of rioting (Metropolitan Police, 2011).

The eventual suppression of the riots is seen to coincide with a significant increase in police numbers, and the question of whether causality can be ascribed is clearly an issue of interest from a policy perspective. Furthermore, if causality is assumed, a natural subsequent question concerns which level of resource would have been sufficient to overcome the disorder.

6.3 Model

The observations of the previous section provide the basis for the main contribution of this work, which is the building of a mathematical model for the London disorder. Modelling the systems involved is a substantially different proposition than doing so for other crimes, partly because principles of offender behaviour are only understood in relatively coarse terms. In addition, the need to incorporate crowd effects distinguishes it from work on the majority of other crimes.

The model proposed here is, accordingly, a macro-level one, concerned essentially with describing the actors in a riot in aggregated terms. That is, the aim of the model is to describe the actions and evolution of the crowd as a whole, for example by giving the total number of (generic) riot participants at a given location at a given time. Nevertheless, the model is designed by considering the stages through which an individual offender is likely to pass in the course of participation. In terms of the typology of models given in Section 5.1.2, therefore, it contains elements of both traditional mathematical and agent-based approaches, though with relatively little emphasis on individual behaviour.

The determination of the particular mechanisms of the model relies on the observation that a number of the findings of Section 6.2 invite analogy with processes arising in other fields of research. These are systems for which models exist and have been well-studied, and so the approach taken here is to adapt, and couple, those models for application to the riots. The resulting model, which will be described in

the sections which follow, is therefore a hybrid of several processes.

The scope of the model is restricted somewhat, in order to remove some sources of significant additional complication. The aim is to model only one generic outbreak of rioting - in effect, one night's disorder - rather than the five which comprised the entire episode. The reason for this is that, in the characteristics which the model seeks to reflect (*i.e.* the generic spatio-temporal evolution), the individual outbreaks are indistinguishable. Although it is likely that the activity on a given night determines the 'initial conditions' for the subsequent night, including the factors which determine these conditions would significantly increase the complexity of the model. In a similar vein, only aggregate offending is considered, rather than particular crime types.

In terms of the aim of the modelling process, the objectives are relatively modest. It is recognised that the reality of a riot is very complex: an extremely large number of factors contribute; many of the interactions involved are non-linear in nature; and their development is highly dependent on the particular setting (*e.g.* the motivation for the offending). To model such a system comprehensively would be infeasible, and may indeed be counter-productive in terms of wider applicability. In accordance with this, no attempt is made to replicate the London disorder exactly or to match the data in a rigorous way; indeed, to do so would imply a false confidence in the quality of the data and risk over-fitting.

Instead, the approach adopted here is a 'generative' one (in the sense outlined by Epstein, 2006), in which a model is sought which can give rise to realistic behaviour and reproduce the general patterns observed for a riot of this type (the 'stylised facts' concerning such an outbreak). Such a model is a proof-of-concept, in this sense, and its ability to generate such patterns suggests that insight might plausibly be gained through analysis of the underlying dynamics, or by examining its behaviour under varying conditions.

6.3.1 Overall structure

Considering the previous sections as a whole, it is apparent that many of the theoretical and empirical observations presented are concerned, in one way or another, with target choice. The construction of the model is therefore approached from that perspective: it seeks to encode the decisions of participants in a way which can be used to derive a representation of the aggregated system. The main concern of the model is therefore the mathematical description of the target choice process.

This has implications for the spatial setting of the model: any choice is made from a certain perspective and, since it is apparent that spatial factors (*e.g.* distance) influence these decisions, it is necessary to consider the locations of participants at the point of choosing. In accordance with this, the main concern of the model is the movement of rioters from their residences to the locations of their offences; *i.e.* the origins of rioters are incorporated throughout. In general, therefore, the key quantities being modelled are the flows of rioters between residential areas and targets.

In order to structure these flows, the model is situated on a discrete spatial system of *zones* (or *centres*). This comprises two distinct entities: residential areas (origins), indexed by i , and targets (destinations), indexed by j . The decision to use a discrete structure is made for a combination of theoretical and practical reasons. On one hand, the majority of data is only available at discrete level, and several assumptions would be required to represent it otherwise. In addition, to include additional detail would be somewhat unnecessary: part of the value of the model lies in the aggregation of effects at an area level. Furthermore, many of the potential outputs of the model are most meaningful when expressed in discrete terms: ‘the number of participants in Brixton’ is of more immediate interest than the continuous distribution of riot participation across south London, for example.

For the purpose of application to London, specific spatial structures are used for both location types: residential areas are taken to be the LSOAs of London, and

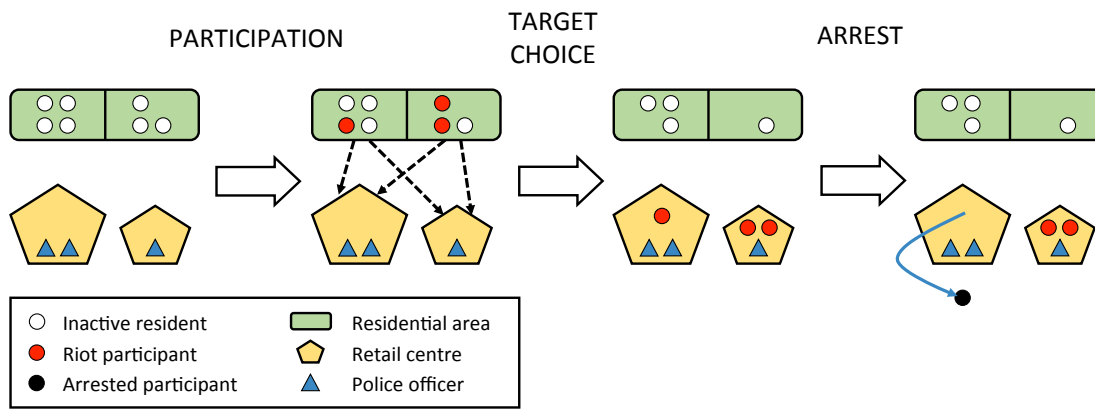


Figure 6.6: The sequence of steps in the composite riot model. In the first stage, the participation model causes residents to become active in the disorder. In the second stage, potential targets are evaluated and rioters move to offend against them. Finally, the interaction between police and rioters causes some to be arrested and removed from the system.

targets are taken to be the set of retail centres (as described in Section 6.2.1). The first of these is chosen for its direct correspondence to UK census data, and the second is chosen because of the observation that the majority of offending in London was directed at retail premises. Although the use of these centres is a substantial simplification, exhaustive representation of all targets would be infeasible, and the set of retail centres is a succinct means of accounting for the dominant mode of offending.

The model itself has three components, corresponding to the three stages through which a generic offender passes in the course of participation. These stages are: the decision to participate, the choice of site at which to offend, and the interaction with police at the location of offending. Figure 6.6 shows a schematic diagram of this sequence, representing one pass through the model of participation. The full model is a dynamic version of this, in which the process is repeatedly iterated. The three stages are modelled by analogy with other systems, each of which reflect the fundamental mechanisms of interest.

In the case of the decision to participate, the analogy draws on the apparent contagious nature of offending, and the model is an epidemiological one. This is motivated

by the qualitative observation that ongoing offending provided encouragement for others to become involved, which is supported by the observed time-course of offending. An analogy is made with a Susceptible-Infected-Removed (SIR) model (see Anderson & May, 1992), which encodes the three stages of infection in a population. The ‘infection’ in this case is participation in the riot, so the progression corresponds to the transition from inactivity to participation and finally to the termination of offending. Of course, neither of those transitions is guaranteed to occur.

The choice of target is modelled by analogy with a retail system; that is, a spatial arrangement of retail centres and sources of spending. Retail systems are a popular application of entropy-maximising spatial interaction models (SIMs), which seek to evaluate the most likely flows of spending from residential areas to retail centres (see Wilson, 1971). The basis of these lies in modelling the preferences of generic individual consumers, in terms of where to spend, as combinations of both spatial and non-spatial factors. Aside from the fact that their spatial settings are identical, the analogy with rioting is clear: in both cases, individual participants make decisions between a number of options, located in space, and the objective is to produce an aggregated description.

Having determined target choice, the final stage of participation is the interaction with police; arrest is the means by which rioters are ‘removed’ from the system. Little previous work has considered the dynamics of interactions between police and offenders in a crowd situation, but one example can be found in the study of civil violence (Epstein, 2002). Borrowing from one aspect of this (agent-based) model provides a means to describe the process of offender arrest, and the interplay between police deployment and crowd disorder.

This three-stage participation model is used to derive an aggregated representation of the system by applying a probabilistic approach to the population as a whole. The quantities contained in the final model are therefore the total numbers of ri-

oters in various parts of the system (*i.e.* partitioned by location and participation state). In this sense, the model is ‘account’-based: the dynamics concern changes in the stocks of actors in various compartments as the system evolves. Table 6.1 summarises, for reference, the main account variables which feature in the model, along with some other key quantities. Each of these will be introduced and defined fully in the description of the model which follows. It should also be noted at this stage that the model is defined for a discrete temporal scheme, with time denoted by t and a uniform time-step δt . All variables listed in Table 6.1 are dynamic, and their evolution will be specified using difference equations for temporal increments of δt .

Variable	Definition
I_i	Inactive residents in residential area i
N_i	Activation rate of new rioters in area i
A_i	Active residents in residential area i
T_{ij}	Rioters from area i targetting site j
R_j	Rioters at retail site j
C_i	Capture rate of rioters from area i
D_j	Demand for police at site j
P_j	Police officers present at site j

Table 6.1: Summary of key variables in the riot model. Each of these represents a count, flow or activation rate for either riot participants or police, and all vary with time.

Although the models can be considered as sequential stages of riot activity, the way in which they are joined incorporates several feedback mechanisms. The sequential component is fairly clear: the participation model provides the input for target choice, and the choice of target determines the dynamics of police interaction at each of the riot sites. In addition, however, the output of the target choice model feeds back to both itself and the participation model, in the sense that the choices of earlier participants influence later ones. The quantity which provides the link between all models, and through which these feedbacks are realised, is attractiveness.

6.3.2 Attractiveness

Any model which concerns the decisions of rational actors requires some concept of the utility of the various choices available. For the model considered here, this is most apparent for the target choice component, in which the merits of offending in various locations are compared. It is also relevant, however, for the decision to participate, which concerns the utility of offending as a whole. In either case, it is necessary to specify the utility of offending at any particular site, j , from the perspective of a generic offender at residential location i .

The determination of such utilities is a necessary step in the specification of any SIM, and is typically formulated using a cost/benefit structure, incorporating factors such as centre size and travel distance. A similar approach is taken here, but with one key innovation: the utility is defined dynamically, and takes into account the current state of the system. The riot system is distinguished by the fact that offender perceptions are affected by the decisions of other rioters and of the police, so that the utility must reflect the state of the crowd in a dynamic way. The cost/benefit structure therefore takes into account three factors: the potential reward at a site; the cost of travelling to it; and the deterrence at a site, which captures the relative numbers of rioters and police.

The *benefit* term, representing potential reward, for site j is taken to be a function of its retail floorspace. It is given by the logarithm of Z_j , which is a non-dimensional measure of its relative size (*e.g.* the ratio of j 's floorspace to the mean across all retail sites). The use of such a form, which reflects the notion of diminishing return to scale, is standard in retail models of this type (Harris & Wilson, 1978; Wilson, 2008). Denoting the benefit of site j as perceived by an individual in i as b_{ij} , therefore:

$$b_{ij} \sim \log Z_j, \quad \forall i. \quad (6.1)$$

Travel *cost* is also incorporated exactly as it is in analogous retail systems, by taking a linear function of the Euclidean distance between the centroids of residential area i and retail site j . Defining this distance as d_{ij} , the corresponding proportional relationship for cost c_{ij} is therefore given by

$$c_{ij} \sim d_{ij}. \quad (6.2)$$

The novel aspect of the utility calculation lies in the notion of *deterrence*, which captures the (dis-)incentive to riot in a certain location due to existing activity. This is calculated by assuming that individuals' discouragement from rioting in a certain location is determined by their perceived probability of arrest if they were to offend there: low perceived chance of capture encourages participation, and *vice versa*. This probability is taken to be a function of the relative numbers of rioters and police officers in a given location, and the particular form used is that proposed by Epstein (2002):

$$\begin{aligned} \mathbf{P}_j^{arrest}(t) &= \mathbf{Prob}(\text{individual in } j \text{ is arrested in } [t, t+1)) \\ &= 1 - \exp\left(-\left\lfloor \frac{P_j(t)}{aR_j(t)} \right\rfloor\right), \end{aligned} \quad (6.3)$$

where P_j is the number of police officers in j , R_j the number of rioters in j , and a is a parameter which represents the number of police officers required, on average, to 'contain' one rioter.

One notable feature of the expression (6.3) is the use of the floor function $\lfloor P_j(t)/aR_j(t) \rfloor$, one of the implications of which is that the probability of arrest is 0 whenever $P_j(t) < aR_j(t)$. Its inclusion has empirical foundation: the Metropolitan Police review of the London disorder (Metropolitan Police, 2012) explicitly states that "decisions were made not to arrest due to the prioritisation of competing demands...specifically, the need to protect emergency services, prevent the spread of further disorder and hold ground until the arrival of more police resources". Accordingly, when the police are 'outnumbered' at a site (*i.e.* $P_j(t) < aR_j(t)$), the

situation is considered to be out of control and the police are unable to make any arrests without the addition of further resources. On the basis that increased probability of arrest corresponds to reduced utility, deterrence is taken to be proportional to this floor function:

$$\text{deterrence} \sim \left\lfloor \frac{P_j(t)}{aR_j(t)} \right\rfloor. \quad (6.4)$$

The full expression for (*benefit - cost*) can now be obtained by combining equations (6.1), (6.2) and (6.4) to obtain the utility of offending at site j , as perceived from site i :

$$w_1 \log Z_j - w_2 d_{ij} - w_3 \left\lfloor \frac{P_j(t)}{aR_j(t)} \right\rfloor, \quad (6.5)$$

where w_1 , w_2 and w_3 are constants.

The motivation in defining this utility is for it to form the basis of a SIM; indeed, its construction mimics that used by Wilson (2008). As derived by Wilson (1970), the *attractiveness* term which features in that model is found by exponentiating the utility term. With a view to this being used as the primary quantity of interest in the model, therefore, the attractiveness of retail centre j , as perceived from residential area i , is denoted W_{ij} and can be written down thus:

$$W_{ij}(t) = Z_j^{\alpha_r} \exp(-\beta_r d_{ij}) \exp \left(- \left\lfloor \frac{\gamma_r P_j(t)}{R_j(t)} \right\rfloor \right). \quad (6.6)$$

The parameters which appear in the expression for W_{ij} - namely α_r , β_r and γ_r (which itself absorbs a) - are parameters to be obtained via calibration with data. These values encode the relative importance of the three factors as determinants of the rioters' behaviour (the subscript r denotes reference to riot participants) and can be varied in order to tune the model as a whole.

6.3.3 Riot participation

The first stage of the dynamic model concerns the initial decision to participate; it is that process which produces a supply of offenders for the subsequent target choice process. Motivated by the hypothesis, consistent with the temporal progression of

the riots, that exposure to nearby disorder had the effect of inciting participation, an epidemiological model is used. In particular, the mechanism is inspired by a compartmental model of the Susceptible-Infected-Removed (SIR) type (Anderson & May, 1992): individuals are assumed to be in one of three states, and the model governs the transition between them.

The participation stage corresponds to the transition from Susceptible to Infected (*i.e.* ‘Inactive’ to ‘Participating’), and so the quantity to be modelled is the rate of transition between these states. This is done independently for each residential area, so that each residential area has an associated independent model, determining the ‘stock’ of riot participants originating from there. The rate itself is found by first specifying the probability that a generic resident of area i will decide to participate during a period of one time unit. This is taken to be

$$\begin{aligned} \mathbf{P}_i^{offend}(t) &= \mathbf{Prob}(\text{individual in } i \text{ chooses to offend in } [t, t+1)) \\ &= \rho_i^\mu \frac{\sum_j W_{ij}(t)}{1 + \sum_j W_{ij}(t)}, \end{aligned} \quad (6.7)$$

where ρ_i is a measure of the deprivation in i and μ an exponent to be calibrated. The deprivation term ρ_i is taken here to be the normalised IMD rank, and is included to reflect the empirical observation that residents of more deprived areas offended disproportionately often. The value of ρ_i is higher in more deprived areas, so that residents are more prone to participate.

The general form of (6.7) is a logistic function of the cumulative attractiveness, $\sum_j W_{ij}$, across all possible riot locations. The cumulative attractiveness is used to reflect the appeal of participation as a whole, though it should be noted that, since distance decay is included in W_{ij} , closer sites will contribute to a greater extent. The logistic function is used to represent threshold-like behaviour, in terms of the point at which rioting becomes appealing; the transition is likely to be localised rather than gradual (*i.e.* a form of ‘tipping point’ exists). Intuitively, the probability (6.7) will be small when the overall attractiveness of potential riot areas is low, whereas,

when the ‘ambient’ level of rioting is high, the probability of offending tends towards ρ_i^μ . From another perspective, if two areas were equally exposed to disorder, greater participation would arise in the more deprived of the two. Social disorganisation theory (Sampson & Groves, 1989) provides a possible explanatory mechanism for this.

This can be translated to the aggregate level for a residential area i by finding an expression for $N_i(t)$, the rate at time t at which individuals choose to participate. It is assumed that decisions are independent between individuals, so that this is given by the product of population size and decision probability:

$$N_i(t) = \eta I_i(t) \rho_i^\mu \frac{\sum_j W_{ij}(t)}{1 + \sum_j W_{ij}(t)}, \quad (6.8)$$

where η is a *transition rate* and $I_i(t)$ the number of inactive individuals resident in area i at time t .

This can be used to formulate equations for the evolution in time of both $I_i(t)$ and $A_i(t)$: respectively, the number of inactive and active rioters whose residence is in a given area i . These, along with their initial conditions ($I_i(0)$ is the residential population of i and $A_i(0)$ a seed of participants, to be chosen) determine the numbers of individuals of each type, in each residential area, at all times. The model is structured in this way so that the composition of rioting groups can be understood in terms of their origins, which is one of the main themes of the work. The equation for I_i also includes an extra term C_i , to be fully defined in Section 6.3.5, for the rate at which participants from i are captured at time t :

$$A_i(t + \delta t) = A_i(t) + \delta t(N_i(t) - C_i(t)) \quad (6.9)$$

$$I_i(t + \delta t) = I_i(t) - \delta t N_i(t) \quad (6.10)$$

6.3.4 Spatial assignment

6.3.4.1 Riot participants

The next stage of the model involves the assignment of active rioters to sites of disorder (*i.e.* retail centres), and this is done using an entropy-maximising SIM. The purpose of models of this type is to estimate the most probable origin-destination flows in a discrete spatial system, given expressions for the attraction exerted by each destination j on each origin i . Such models are also constrained by expressing, for example, the total out-flow from each origin and/or the total in-flow at each destination. A full review of such models, their derivation and applications is given by Wilson (1971).

The riot model, as constructed, corresponds to this type: attractiveness is given by the terms W_{ij} and the total out-flow at each residential area i is given, at any time, by A_i , the number of active rioters originating there. Before applying the model, however, a small modification is made to the attractiveness term: rather than using W_{ij} itself, a moving average over a number of time-steps is used. This is done for several reasons: to account for factors such as travel time on the part of rioters; to represent ‘lag’ in the spread of information through the system; and to dampen the effect of sudden fluctuations in attractiveness. Therefore, the values used to determine spatial assignment at a given time, referred to as *effective attractiveness* and denoted W_{ij}^e , are the average values of W_{ij} over the L_r most recent time-steps in the discrete temporal scheme (when $t < (L_r - 1)\delta t$, the summation is padded with the $t = 0$ value):

$$W_{ij}^e(t) = \frac{1}{L_r} \sum_{l=0}^{L_r-1} W_{ij}(t - l\delta t). \quad (6.11)$$

Given the expressions for W_{ij}^e and A_i , the application of the model is straightforward; following the standard entropy-maximising derivation given by Wilson (2008), it can be shown that T_{ij} , the number of rioters from i who are participating in disorder in

j , is given by:

$$T_{ij}(t) = \frac{A_i(t)W_{ij}^e(t)}{\sum_k W_{ik}^e(t)}. \quad (6.12)$$

The expression can be summarised in intuitive terms: the proportion of flow assigned to a given centre j is given by the proportion of the total attractiveness, across all centres, which is attributable to j . As an aside, it can also be noted that (6.12) can be derived alternatively by interpreting (6.5) as the utility term in a conditional logit model (McFadden, 1984, as used in Chapter 4): a duality exists between the two frameworks. In either case, summing over residential areas i yields the total number of rioters, R_j , at a given retail centre, j :

$$R_j(t) = \sum_i \frac{A_i(t)W_{ij}^e(t)}{\sum_k W_{ik}^e(t)}. \quad (6.13)$$

It should be noted here that each time unit is therefore defined implicitly as the mean time taken for each participant to travel from a home location to a chosen target.

6.3.4.2 Police resources

The assignment of police resources to areas of disorder is also realised via a SIM, as for riot participants; there are, however, noteworthy differences. First, police units are not considered to have a ‘home’ location and are active and situated at potential sites of disorder at all times. In addition, the response lag L_p is also different to that of rioters (and intended to be higher), reflecting the delay in learning of the plans and movements of rioters, and conferring upon the rioters a degree of ‘first-mover advantage’.

The main difference for police, however, is in the attractiveness function, the equivalent of which for police is the *demand* for officers at a given site. Following a similar argument to that which was used to derive (6.5) for riot participants, it is assumed that the *benefit - cost* of police follows:

$$\widetilde{w}_1 \log Z_j + \widetilde{w}_2 R_j(t), \quad (6.14)$$

for some parameters \widetilde{w}_1 and \widetilde{w}_2 . This expression (6.14) includes no spatial decay term, reflecting the fact that the police do not prioritise incidents on the basis of proximity (HMIC, 2011) and can travel to incidents rapidly. In addition, the second term is a function of rioter numbers only: given that their aim is to eliminate all disorder, the number of police already at a site is likely to be immaterial to the police. Analogously to the attractiveness derived in (6.6), the function D_j representing police demand at site j is therefore:

$$D_j = Z_j^{\alpha_p} \exp(\gamma_p R_j(t)), \quad (6.15)$$

where α_p and γ_p are, as before, parameters to be calibrated which encode the relative importance of the two factors. As with attractiveness, temporal lags are accounted for by deriving the time-smoothed *effective requirement*:

$$D_j^e(t) = \frac{1}{L_p} \sum_{l=0}^{L_p-1} D_j(t - l\delta t). \quad (6.16)$$

where L_p is the lag experienced by police (comparison of L_p and L_r gives the relative responsiveness of the two groups).

Equation (6.16) provides the necessary input for a SIM to estimate police assignment. In contrast to the case for rioters, it is assumed that the total number of police officers in the system is a constant value, P^{total} , determined exogenously; that is, there is no analogous infection-like process. It is therefore simple to apply the spatial interaction expression to find the total number of police officers in location j at time t :

$$P_j(t) = P^{total} \frac{D_j^e(t)}{\sum_k D_k^e(t)}. \quad (6.17)$$

6.3.5 Interaction between police and rioters

The final element of the model concerns the interaction between police and riot participants at the locations of disorder. This takes the form of a process of arrest, through which riot participants are removed from the system (this corresponds to

the Infected-Removed transition in the epidemiological framework). Again, the rate of transition is estimated on the basis of the probability of arrest of a generic rioter in location j . An expression for this probability has already been given in (6.3) - it is the primary determinant of deterrence in the original attractiveness calculation - and so the same formulation can immediately be applied.

For reasons explained previously, riot participants are accounted for primarily on the basis of their home location, and so their removal is calculated separately for each residential area i : this gives the arrest rate term C_i which appears in (6.9). It is found by multiplying, for each riot location j , the number of participants from i who appear there and their probability of arrest. This is then summed to give the aggregate rate:

$$C_i(t) = \tau \sum_j T_{ij}(t) (1 - \exp \left(- \left\lfloor \frac{P_j(t)}{R_j(t)} \right\rfloor \right)). \quad (6.18)$$

where τ is an *arrest rate* parameter.

6.3.6 Model integration

In order to produce the full integrated model, the various sub-models described above are implemented in sequence, and numerical simulations carried out by iterating across the relevant equations. At each time-step, the first task is the evaluation of the core quantity of attractiveness, W_{ij} , upon which several of the following calculations are based. This acts as the input for the participation model (equations 6.8 and 6.9), which gives the number of rioters active in the system, A_i . In the next stage, these rioters and the available police resources are distributed by applying equations (6.13) and (6.17) simultaneously, giving the numbers of each type of actor at each retail centre. Finally, the number of rioters arrested is calculated according to (6.18) and these are removed from the population.

As remarked previously, this mechanism incorporates a number of feedbacks, with attractiveness providing the key link between processes. With the relevant equations now specified, these relationships can be sketched in terms of the particular vari-

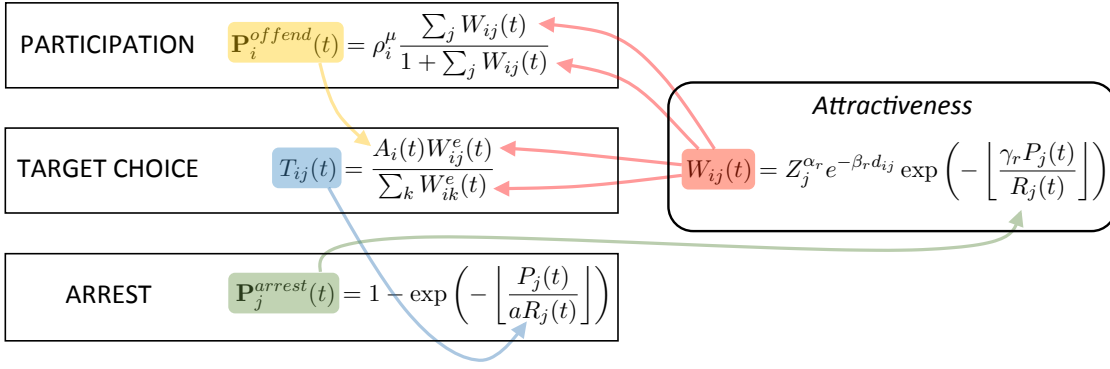


Figure 6.7: Feedback relationships in the composite model. The diagram shows the key expression from each of the three stages of the model, together with the unifying quantity of attractiveness. The arrows indicate that the latter term is equal to, or determined primarily by, the former.

ables which are influenced. Figure 6.7 shows the key equations from each stage of the model, together with attractiveness, with the main inter-connections identified.

The yellow and blue arrows indicate that, for both participation and target choice models, the resulting value is used to determine the input for the next stage in the sequence. The probability of arrest, calculated in the final stage, has a dual role: it governs the removal of rioters and is also the mechanism by which crowd effects feed back into the model via attractiveness, as indicated by the green arrow. The role played by attractiveness in the first two stages, shown by the red arrows, completes the feedback cycle via the influence on the behaviour of rioters at later stages.

6.4 Results

As a result of the desire to capture the complete process of riot participation, comprising several stages, the model proposed in the previous section is a complex one. Its intricacy is such that formal mathematical analysis of the full model is infeasible, beyond trivial observation. The aim of this work, however, was not to propose an analytically-tractable model; rather, the approach was intended to be ‘generative’, in the sense defined by Epstein (2006). According to this principle, detailed understanding of the mechanics of models is not a primary concern, and the main focus instead relates to their sufficiency to reflect observed phenomena. In this sense, the

approach has much in common with agent-based modelling (indeed, the model could be translated into such a form).

In line with this, the model is examined primarily through numerical simulation, and a number of these will be presented in this section. The first acts as a demonstration of the validity of the model by presenting a configuration (*i.e.* a combination of parameters and initial conditions) which is capable of reproducing the stylised characteristics of the London riots outlined in Section 6.2. With this established, the remaining analysis examines two issues which relate closely to key policy questions arising from the riots: the susceptibility of individual sites, and the effect of varying police resources. These are intended to act as proofs of concept for the use of a model such as this to provide quantitative insight in a policy context.

6.4.1 Numerical simulations

The numerical simulations referred to throughout this section were performed in Python, using step-wise iteration of the equations as described in Section 6.3.6. The discrete time-step used throughout was $\delta t = \frac{1}{60}$, and all simulations were run for a total of $t_{max} = 10$ units. Although these temporal units are intended to be non-dimensional, their scale is such that one run of the model could be regarded as corresponding to a 10-hour episode of rioting (corresponding to one night's disorder), with the system being updated each minute.

The initiation of the model involves setting values for both initial conditions and model parameters. Those initial conditions which are common to all scenarios - the size of retail centres, and the population and IMD score of each LSOA - were first read from the corresponding datasets and associated with each of the relevant spatial units. The remainder - those which are the subject of investigation and therefore allowed to vary between simulations - were then set as desired: the 'seed' of active rioters at initiation, and the number of police officers, for example. Together with a given set of parameter values, this was sufficient to simulate one run of the system.

6.4.2 Demonstration case

The objective of the first stage of analysis was to establish whether the model was capable of generating patterns of sufficient similarity to the observed data that it could be regarded as a plausible representation of the London riots. Since the output of the model is determined, to a large extent, by the parameter values used, this task is equivalent to examining whether a configuration can be found which gives rise to a simulation with the desired characteristics, and therefore reduces essentially to a search of parameter space.

For the purpose of this parameter search, simulations were performed using a simple choice of initial conditions, in order to represent a generic episode of disorder. The system was initialised with a seed of 100 active riot participants, distributed across residential areas in proportion to their population and allocated to retail sites in proportion to the static component of attractiveness, $Z^{\alpha_r} \exp(-\beta_r d_{ij})$. The number of police officers, P^{total} , was taken to be 5,000 (the approximate number deployed across the first 3 days of disorder in London), and, similarly, these were distributed across retail sites in proportion to the static component of demand, Z^{α_p} .

Given that the aim of this process was to produce results which were, in some sense, ‘plausible’, it is necessary to specify the terms in which this plausibility was evaluated. This highlights one of the difficulties with this aspect of the analysis: the question of how to assess a given model output is, in general, a non-trivial one, since there is no single quantity which encapsulates all relevant characteristics of an episode of disorder. Both the journey-to-crime distribution and temporal progression, for example, were identified as notable features of the observed data and therefore ought to be reflected, yet are clearly distinct properties.

The approach taken here is to identify a number of key characteristics for which simulated and observed data can be compared quantitatively; an output is then considered plausible if all measures lie within a certain tolerance. For example,

Parameter	α_r	β	γ_r	α_p	γ_p	η	k	τ	L_r	L_p	δt
Value	0.6	0.5	0.11	0.65	0.012	0.006	6	0.75	30	60	0.0143

Table 6.2: Parameters used in demonstration case simulation

the distribution of flow-weighted travel distances (*i.e.* the final flows T_{ij} occurring over each distance d_{ij}) can be measured, in terms of Kolmogorov-Smirnov distance, against the observed journey-to-crime distance. Though this is somewhat imprecise, this approach is commensurate with the aim of the work: no attempt is made to ‘match’ the observed data quantitatively, and the objective is simply to produce qualitative similarity.

In addition to travel distance, the other observables computed in each case were: the distribution of riot magnitude across all sites; the number of riot sites where the level of offending was of an order higher than the mean level; and the growth of total rioter numbers over time. Each of these ensures that runs which exhibit blow-up, or uniform escalation across all retail sites, are excluded from the analysis.

The parameter space to be searched is relatively high-dimensional, so that exhaustive search is impractical, and the process was therefore carried out iteratively. In the first instance, parameter combinations were taken from a coarse sampling of the space and outputs evaluated in terms of the diagnostics described above. Using these, a smaller region of parameter space was identified for which all observations were of similar character to the riot data (in the sense that their relative difference was within a certain tolerance). The process was then repeated for the smaller parameter space, using a lower tolerance, and several further similar iterations followed.

In order to focus discussion, and to provide a base case for the remaining analysis, one particular set of parameter values was chosen, and these are shown in Table 6.2. Although the corresponding simulation output does not replicate the observed data exactly, it shows qualitative similarity in several respects.

The journey-to-crime in simulation outputs is considered first. Figure 6.8 shows the weighted flow distribution, which is approximately exponential in form, as was the case for the observed data (Figure 6.1). This agreement is one of the properties which led to the selection of this case. Nevertheless, some disparity does remain between the observed and simulated distributions: this can be ascribed to the influence of many factors which are not incorporated in the model, and is a consequence of the desire for a simplified model.

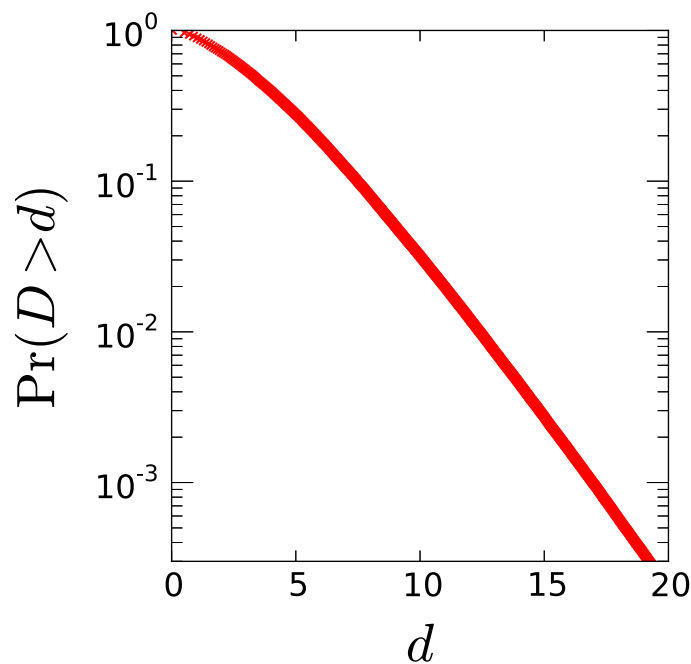


Figure 6.8: The distribution of flows T_{ij} for each inter-centre distance for the output of the demonstration case simulation.

Figure 6.9 shows the particular case of Brixton; specifically, it shows the origin locations of simulated disorder in Brixton. This shows a similar distance decay form to that seen for the observed data in Figure 6.2, although it is notable that the relationship is not uniform in space. This can be accounted for by two factors: the influence of deprivation on offending, and the fact that residential areas are affected differently by the ‘competition’ between multiple retail sites.

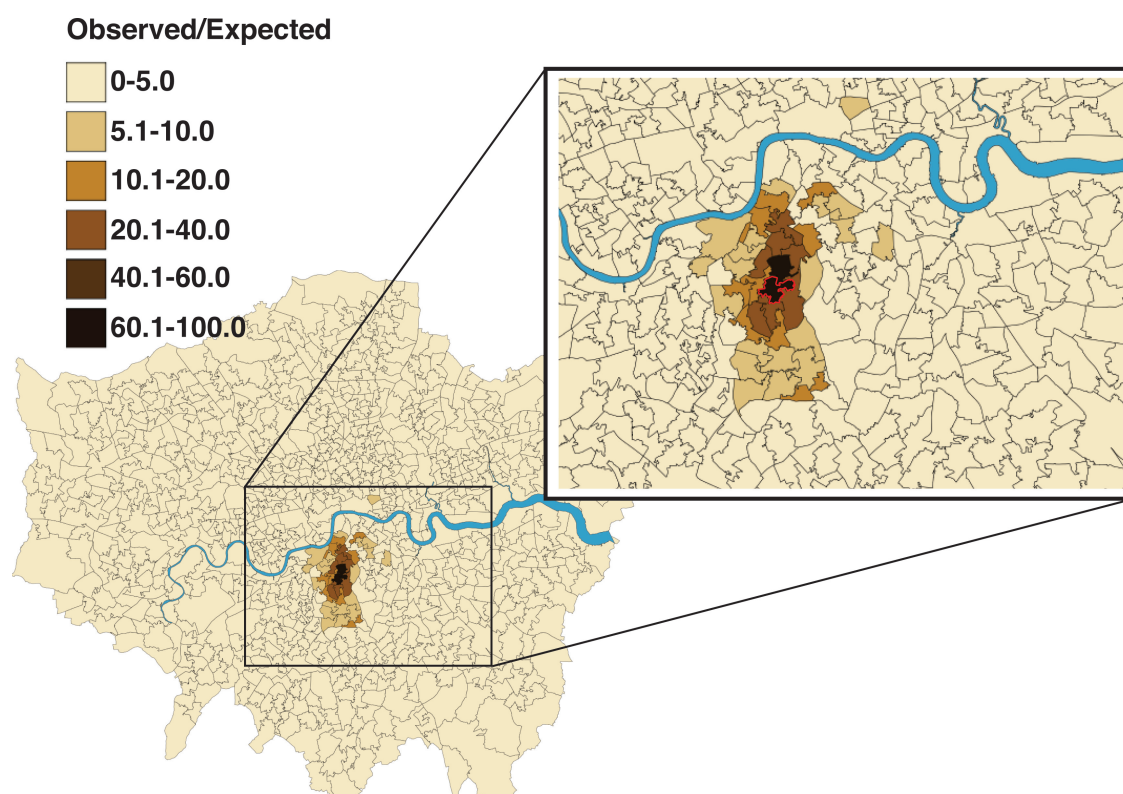


Figure 6.9: LSOAs coloured according to the proportion of residents offending in Brixton (outlined in red) in the output of the demonstration case simulation.

One further issue of interest is the aggregate distribution of offender residences, and this is shown for observed and simulated cases in Figure 6.10. Comparing the two reveals both agreement and disagreement: on one hand, almost all high-participation areas are identified by the model, including the known prevalence in south London. The simulated data, however, predicts a number of additional sites and suggests a different relative importance; in particular, areas of north-east London are found to have very high participation. It is likely that this is an artefact of those areas' high deprivation and their proximity to the large retail centres of central London: this causes high participation in the model but, for unknown reasons, the effect was not manifested in the riots themselves. Overall, the agreement is fairly good: 26 of the 33 boroughs show modelled rioter percentages in the same or adjacent band to that observed in the data.

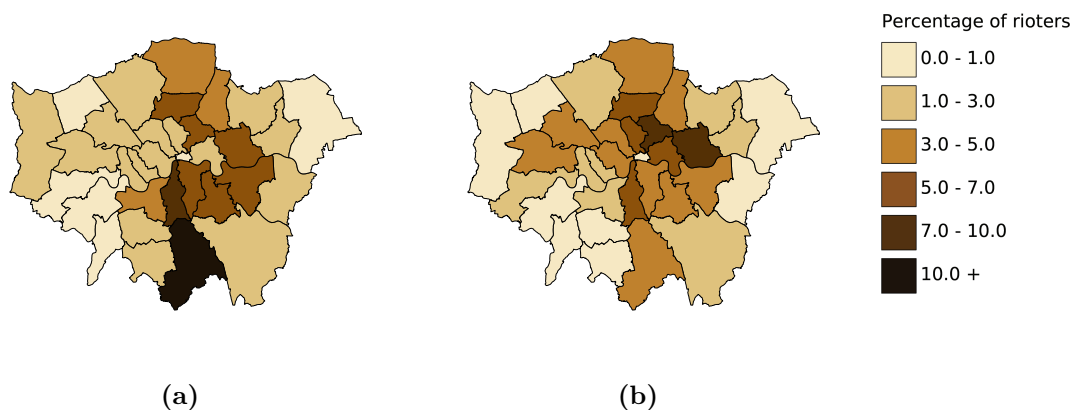


Figure 6.10: Variation in participation levels across London, for a) observed, and b) simulated data, shown as borough-level choropleth maps.

6.4.3 Susceptibility of sites

The simulations used to establish the base case were performed using generic initial conditions, with offenders spread evenly across the city. This is a significant simplification, however: in reality, riots tend to begin with small pockets of disorder, and indeed this was the case in London. The next step in examining the behaviour of the model is therefore to consider initial conditions of this type and the evolution under this more realistic premise.

The way in which this could be done - the array of possible initialisations - is almost limitless; however, using one of the key policy concerns as a motivating question provides focus for the task. A significant consideration in the issue of why riots occurred in some places and not others is whether locations had some inherent susceptibility to disorder or whether the offending there was simply due to the particular circumstances of the 2011 events. That is fundamentally a question of robustness; that is, the extent to which changes in the initialisation of the model affect the locations of disorder.

These issues can be examined simultaneously by comparing the results of simulations in which disorder was initialised in a small number of nearby locations. For four of the worst-affected sites in London - Brixton, Clapham Junction, Croydon

and Ealing - a series of simulations were run in which the system was initialised with small disturbances at both the site in question and one of its closest geographical neighbours. In essence, a series of contests are run between close neighbours, after which it can be seen whether one consistently experiences disorder. The parameters used in all simulations were those of the demonstration case, as established above and listed in Table 6.2.

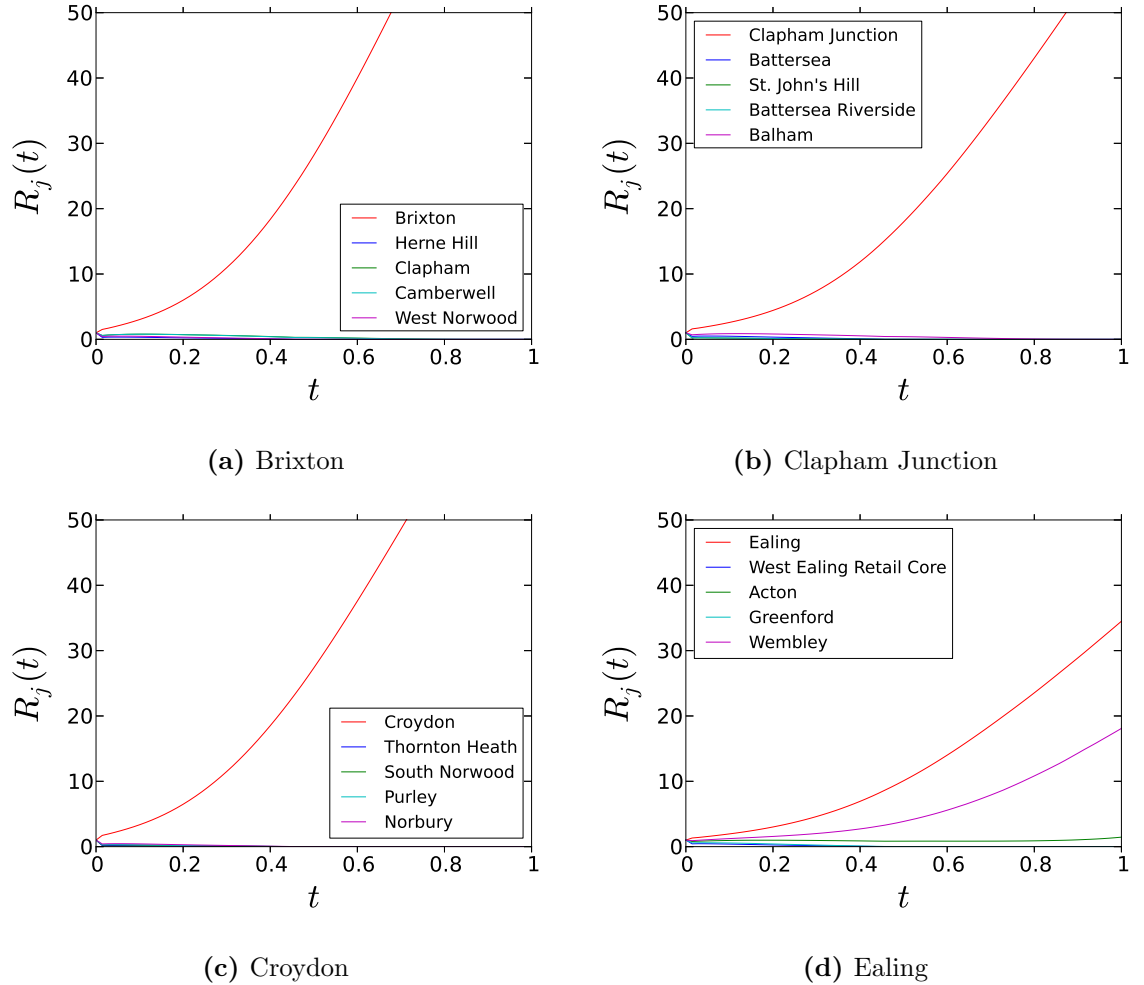


Figure 6.11: Susceptibility amongst local groups of retail sites. Each panel represents the outcome of four simulations; in each, disorder was initialised at the named site and one of its four closest neighbours, and the number of riot participants at each site, R_j , monitored over time. The red line in each case represents the average over the four simulations in which that site was perturbed.

Figure 6.11 shows the results of such simulations, in which an initial disturbance of one rioter is used (*i.e.* $R_{j_1}(0) = R_{j_2}(0) = 1$ for each pair of simulated sites j_1 and j_2). For each of the four sites of interest, the red series represents an average over

the four simulations in which it featured, with the others presented for comparison. In three cases, there is a rapid escalation in disorder at the site which was affected in the 2011 riots, while disorder decays to negligible levels for all neighbouring sites examined. The exception to this is the case of Ealing, for which disorder also grows at Wembley; Ealing is still, however, the dominant site.

These results suggest that, insofar as the model describes a generic episode of rioting in London, the sites in question have particular susceptibility to disorder. That one site should dominate is unsurprising, since this is a general property of spatial interaction models, accentuated by the crowd effects present here. However, the fact that the dominant site is the same in all cases suggests that those sites are rendered particularly susceptible by characteristics specific to them. The reasons for this can be inferred from the structure of the model: the sites' proximity to populous areas of high deprivation, a relative lack of local competition from other prospective targets, and low prioritisation by police due to relatively small size.

6.4.4 Police resources and response

The remaining policy questions which motivate this work relate to policing: specifically, the number of officers available during the riots and the mode of their response. In this context, the utility of the model lies in its ability to explore counterfactual scenarios, corresponding to policy interventions, in a quantitative way. This is achieved by implementing variations to the demonstration case model and assessing the effect on the outcome of the riots.

From a policing perspective, the objective when responding to a riot is to minimise, in some sense, the magnitude of the disorder. In order to quantify this, and therefore to compare realisations of the system meaningfully, it is necessary to define this magnitude in terms of the output of the model. Since many of the terms in which the negative impact of disorder is typically measured - such as the cost of damage or extent of personal injury - are not quantified by the model, it is necessary

to propose an alternative. The measurement used will be the *severity* of an episode, denoted by S and defined as the cumulative extent to which police are outnumbered by rioters across all sites and all time:

$$S = \sum_j \sum_t^{t_{max}} \frac{R_j(t)}{P_j(t)}. \quad (6.19)$$

This definition is complementary to that of the probability of arrest (see equation 6.3), in that it equates the extent to which a situation is ‘out of control’ with the imbalance between police and rioter numbers. This loss of control is assumed to determine the extent to which riot participants are able to cause damage, and therefore severity is a proxy measure of the magnitude of a riot. It can therefore be used as a non-dimensional statistic by which realisations of the model can be compared.

The first parameter to be varied was the total volume of police resources, P^{total} , with all other parameters as in Table 6.2. Figure 6.12a shows the variation in severity across a range of values which correspond approximately to the numbers of officers deployed in London (see Figure 6.5). The results indicate, as expected, that higher officer numbers result in lower severity, but the relationship is not linear. After decreasing sharply for relatively low values of P^{total} , the rate of change of S gradually decreases, reflecting diminishing returns to scale as police numbers increase.

These results agree with intuition: when police deployment is low (particularly for $P^{total} < 3,000$), addition of resources has a relatively large effect and disrupts the disorder accordingly. On the other hand, the influence of increasing P^{total} ultimately diminishes, as the marginal ‘cost’ required to decrease S becomes greater. The decrease appears to be relatively steady for values of P^{total} greater than approximately 8,000, suggesting that gains beyond this level are relatively incremental.

This result implies that an intermediate level of police resource, between the 6,000 and 16,000 officers deployed on the final two days of rioting, may have been sufficient

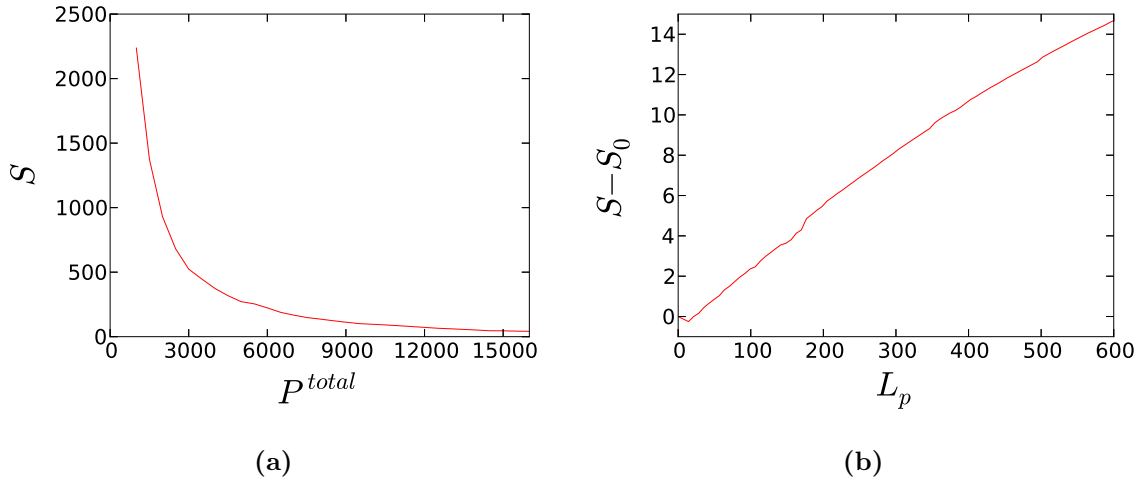


Figure 6.12: Variation of severity, S , as parameters governing police response are varied: a) shows a non-linear decrease in overall severity as overall police numbers, P^{total} are varied; b) shows the increase as police response lag L_p is varied, with the values plotted relative to the $L_p = 0$ case, denoted S_0 .

to contain the London disorder. This is, of course, merely a demonstration of the kind of estimate that a model such as this can be used to provide: the limitations of this model are such that the real-world inferences cannot be made with confidence. Framed as a proof-of-concept, however, evaluation of this type shows the way in which work of this type can be used to quantify policy recommendations.

The same is true for the investigation of the response speed of police, with the results shown in Figure 6.12b. The different simulations correspond to variation in the value of police response lag, L_p , and the vertical axis shows the increase in severity relative to the $L_p = 0$ case (denoted S_0). After a noisy stage at small values, S is shown to increase with lag, so that, as expected, slower police response results in a larger riot.

Although the increase in S is small, relative to its absolute value, it should be noted that these simulations are set up with parameters such that a certain level of severity is assumed. The demonstration case, upon which these simulations are based, represents a substantial riot, and no police strategy could fully eliminate disorder: any changes in severity are variations around a level which has been im-

plied *a priori* by other factors, such as the number of police officers and influence of deprivation.

6.5 Discussion

The aim of this chapter was to explore the use of modelling in a setting which contrasts with that of burglary, around which much of the preceding work was focussed. Large-scale civil disorder, or rioting, is such a crime, and was explored here using the particular case study of the London riots of August 2011. Events of that type are sufficiently rare, and dependent on particular circumstances, that narrowing the scope to one specific example increases significantly the tractability of the problem; more practically, the London episode is one of the few for which it is possible to obtain granular data.

Modelling episodes of rioting presents a number of distinct challenges to those that arise for burglary. The differences can be summarised as being issues of scale, both in terms of the mechanisms which a model seeks to encode and the questions which it seeks to address. On one hand, theory governing the actions of offenders is specified much more coarsely in the case of riot: little is known about the influence of environmental factors, for example, and few hypotheses concern specific target selection. Whereas burglary analysis might consider decision processes at the level of streets or individual properties, riots are more commonly described at the area level.

This reasoning also extends to the questions such a model might be expected to address. Policy issues arising from riots typically relate to large-scale issues - such as city-level preparedness, the magnitude of outbreaks, and the overall strategy of the police - and potential interventions are likely to be of that form. To contribute in that way, therefore, models should be constructed to offer meaningful results in those terms.

The most significant challenge from a modelling perspective, however, concerns the

fact that it is necessary to consider interactions between offenders. Although burglary and riot modelling both aim to describe aggregate patterns, in the case of burglary these are simply the accumulation of individual acts; for riots, on the other hand, significant crowd effects are present. Both the decision to offend and the choice of target are influenced by the activities of others, since riot activity is generally more productive when carried out in concert.

In addition to this, it is clear that riot participation is a multi-stage process, and a model must consider more than simply target choice. Where offenders originate, for example, has a material effect, since it determines the number of offenders in proximity of a potential target. This reasoning, along with the crowd effects noted above, implies that any model must be a complex one.

It is infeasible to represent mechanisms as involved as these in a simple mathematical formulation, and so the approach used here requires compromise in terms of the analytical tractability of the model. Although defined using mathematical expressions, the model proposed has much in common with an agent-based model: arbitrary levels of complexity are added as desired, and the final model is an *ad hoc* combination of existing models, adapted from other fields. The extent of feedback and non-linearity within the system means that it cannot be analysed formally.

The goal of the modelling process was set in accordance with this approach. The aim was to propose a model capable of generating patterns of offending broadly similar to those seen in London, with this serving as a partial demonstration of its validity. No attempt is made to establish the necessity or sufficiency of model components, or to replicate the London data exactly, since the model is not intended to be a comprehensive one. Instead, the model is intended to be a proof-of-concept which can be used to demonstrate the utility of the approach in a policy context.

6.5.1 Findings

Before formulating the model itself, some initial analysis of riot data was performed in order to establish a number of fundamental characteristics of the London riots, against which model outputs could be compared. The reason for this analysis was simply to provide motivation for the latter work, and it was therefore done in a relatively coarse way; nevertheless, a number of signatures could be identified.

In particular, comparison of the home and offence locations of riot suspects revealed a clear distance decay trend, and a positive relationship was also observed at the area level between deprivation and the tendency of residents to offend. These findings are supported by more detailed similar work elsewhere, and are consistent with various criminological theories, such as the rational choice perspective and social disorganisation theory. Taken together with others, these signatures are a distillation of the spatio-temporal character of the London riots.

The first part of the analysis concerned the identification of a ‘demonstration case’: a simulation which showed approximate agreement with the key characteristics of the observed data. A suitable parameter configuration was found, and this was used as the basis for the analysis which followed. Although agreement with the prescribed characteristics was approximate only, the inconsistency can be attributed to factors lost in the simplification of the model.

This discrepancy is reflective of a general difficulty when modelling rare events such as these. In order for a model to have value, it must apply, to some extent, in a more general context; however, the fact that such events have substantial dependence on particular circumstances means that quantitative agreement is therefore difficult to achieve. In this case, it is assumed that this could be resolved by increasing the sophistication, and specificity, of the model in order for it to more realistically represent a specific scenario.

The addition of such sophistication was not, however, the focus of this work; instead, the demonstration case was used to explore possible real-world applications of the model. The first of these concerned the susceptibility of particular areas to rioting, and this was explored by initialising small, localised pockets of riot activity. It was found that, in general, the explosion of rioting in some areas was robust to variation in the initial conditions, from which it was inferred that these sites were subject to some inherent risk.

Again, this highlights more general issues with the approach used. Although the finding that certain sites are prone to disorder is of clear value, it reveals nothing about which of their properties are responsible for the heightened risk. The number of possible causes is limited, of course, to those which are encoded in the model, but their relative contributions are unknown. This is problematic if more general conclusions are to be drawn, and highlights a shortcoming of the generative approach: insight could only be gained by brute-force simulation. In this scenario, this may be a necessary cost of achieving the required model complexity.

The final examination of the model concerned the police response to disorder. The model was tested by varying the number of police officers available and their speed of response, with model runs then compared using a proxy measure of riot magnitude. In both cases, the trends were as expected: greater officer numbers and faster response reduce the severity of riots. These findings are unsurprising, but their value lies in the fact that they are explicitly quantitative: they show diminishing returns to scale, and the levels required to constrain the riots to a prescribed extent can be found. Little weight should be attached to the precise figures, for reasons outlined above, but this demonstrates what is possible with a model of this type.

6.5.2 Future work

One consequence of the proof-of-concept approach used here is that there is significant scope for further work. This concerns both refinement of the model itself and

its applicability in a practical context, in line with the ultimate goal of producing a tool which can be used to perform *in silico* scenario testing.

Several aspects of the model could be made more sophisticated, in light of what is known qualitatively about the riots. It has been widely speculated, for example, that social networks (in particular, electronic social media) played a significant role in the spread of riot participation and the recruitment of offenders. This consideration is absent from the simple model of contagion used here, but could be implemented in a future iteration. In particular, this could be achieved by introducing long-range communication of the type facilitated by social media, rather than the local formulation currently used.

Another example of a potential long-distance interaction concerns the use of non-pedestrian transport networks in the journey to offence. This would, of course, undermine the assumption of simple distance decay in the model, and substantially complicate the target choice decision process. The transport network was omitted from the present work partly for reasons of parsimony, but also because of the difficulty in understanding its structure during the riot period. The very fact that disorder was occurring resulted in transport disruption, so that the true structure would be difficult to obtain. Nevertheless, this remains an avenue for development, if suitable data can be obtained.

As well as the model itself, there is also scope for improvement in the way that it is initialised. In order to simulate a more realistic riot, it is necessary to find a more accurate representation of the initial conditions than the arbitrary local perturbations that are used here. This also relates closely to another point: the extension of the model to several days' disorder, rather than the one considered here. It is likely that the disorder on any given day is influenced by what has gone before (targets may have been exhausted, or knowledge may have been gained about police deployment) and so such a multi-day model would most likely be implemented by

using the output from any day to initialise the next.

One further shortcoming of the model, at present, is in the activity of police. Resources are deployed using a naive mechanism which simply calculates relative requirement; in a real-world situation, such decisions would be taken centrally by a command team. Such central command, and the strategic aspect of their actions, would be challenging to achieve in a model of this type. In one possible development, however, this could be removed from the model and re-cast as a dynamic input; that is, a user could control officer deployment.

The use of the model in this way, with the user able to test resource allocation strategies, is a natural way for this work to be applied in a policy context. Such a possibility is distinguished from other ‘table-top’ approaches by the quantitative and dynamic underlying model, and so could be considered to be a more rigorous means of examining possible approaches. The benefits of this are clear: for a rare event such as a large-scale riot, for which it is infeasible to carry out real-world practice exercises, modelling offers an opportunity for limitless scenario-testing with quantitative output. A basic version of such a tool has already been implemented, and will be developed further.

Chapter 7

Summary and discussion

The aim of this thesis has been to explore the use of techniques from complexity science in the study of crime and security, drawing on examples from a number of particular types of crimes. In particular, the main issue of interest has been the spatio-temporal character of criminal behaviour, and the various means by which this could be both analysed and modelled. Each of the examples considered in the preceding chapters concerned the understanding or prediction of patterns in time and space, though these took various forms and encompassed a range of criminal behaviour. This chapter will begin by summarising the work, before moving on to consider common themes which arose throughout the thesis, and opportunities for further development.

7.1 Summary

The work in this thesis has been motivated by the desire to contribute to the understanding of crime, and its spatio-temporal distribution in particular. This aspect of crime is a central concern in both practice and theory, and is also particularly amenable to quantitative analysis. Indeed, a substantial volume of recent research has examined crime in this way, with notable results: analysis has improved the understanding of clustering, for example, and the success of models in reproducing common patterns suggests that such approaches may be used for prediction. Scrutiny of these techniques, however, reveals a number of opportunities for for-

malisation or increased sophistication, particularly with regard to the way in which environmental factors, and individual-level theories of behaviour, are incorporated in analysis and modelling. These issues have been addressed through the use of techniques from complexity science.

The contributions of the thesis are both practical and theoretical. Several of the findings constitute improvements in the understanding of the spatio-temporal distribution of crime: these include the identification of signatures in clustered patterns of offending, and examination of the role of the street network in shaping urban crime. Each of these is of clear practical use, and also contributes to the theoretical understanding of offender behaviour. The modelling work presented, concerning both burglary and civil disorder, is also of potential use in both day-to-day policing and long-term policymaking. In addition to this, these topics have also provided a platform from which a technical contribution could be made, in relation to the analysis and modelling of social systems. A number of substantial challenges arose during the various components, which required methodological innovation and which have led to findings that are applicable in a wider context. Many of these relate to networks, and the complications which arise in a spatio-temporal context: null models for statistical analysis have been developed, new network metrics have been proposed, and new insights into the behaviour of stochastic dynamical systems on networks have been gained.

Chapter 2 demonstrated the use of a network-based framework as a means of characterising the spatio-temporal clustering patterns present in sets of crime events. Networks derived from data for two crime types - burglary and maritime piracy - were examined for the occurrence of small sub-structures, referred to as ‘motifs’ and ‘chains’, which have real-world interpretation as linked groups of events. Significant configurations were identified for both crime types; furthermore, since the motifs identified in each case were different, it was shown that the method could be used to discriminate between clustering patterns which existing tests would find to be equiv-

alent. The primary technical challenge of the chapter concerned the adaptation of network analysis techniques for application to the particular class of network considered (the spatio-temporal nature of which represents a significant complication). Suitable null models were constructed using a novel computational approach, based on simulated annealing, and the method developed could be applied to analyse a number of other properties of similar datasets.

Chapter 3 was more involved, in terms of its criminological motivation, and concerned the relationship between urban form - in particular, the street network - and crime risk. The work addressed a number of shortcomings of previous research by taking a formal approach to the analysis of street network structure. The network metric 'betweenness' - which reflects travel patterns and therefore corresponds closely to criminological theory - was considered in particular depth, and a statistical model was used to examine its relationship with burglary risk. This relationship was found to be highly significant, which provides support for criminological theory. The findings are also likely to be of practical use: the fact that betweenness is a static, objective and highly granular measurement suggests that it should be of particular value in the assessment of risk.

In Chapter 4, analysis of street network effects on residential crime was extended to dynamic effects; specifically, their role in (near-)repeat victimisation. Although that phenomenon has been observed and studied widely, it had not previously been investigated explicitly in the context of the network; furthermore, few studies have investigated the existence of directionality in the spread of risk. In this chapter, it was shown that not only can space-time clustering be expressed in terms of the street network, but also that the risk of follow-up victimisation does not propagate uniformly in all directions. By representing the selection of a follow-up target as a choice problem, it was shown that the choice was predicted, to an extent, by a street segment-level variable designed to reflect the awareness space of a returning offender. This implies that, for the purpose of crime prevention, streets can be

prioritised on the basis of their position in the network. The variable used in the analysis, commonality, is a novel network metric which extends the principle of betweenness to a dyadic relationship. As a quantification of the extent to which any two network elements feature together in flows through a network, it has a number of potential applications in network science.

Motivated by the findings of Chapters 3 and 4 that street networks play a significant role in shaping patterns of urban crime, the aim of Chapter 5 was to propose a model of residential burglary in which the street network was incorporated throughout. The model proposed was based on the concept of risk diffusion, as in similar models elsewhere, but with the network providing the substrate for the process. In addition, the model differed from previous work in its treatment of awareness space: rather than modelling it directly, this was estimated using network properties (such as betweenness). In terms of modelling, this was significant: the observation that awareness could be regarded as static, on the time-scale of the model, meant that it was not necessary to model offender movements directly (via an agent-based approach, for example). The resulting model was therefore tractable analytically, with one of the key findings being the relationship between the form of crime patterns and the structure of the street network. In particular, it was shown that areas of common risk can be reconciled with community structure within streets (which reflects, for example, the ‘insularity’ of sections of the network).

The aim of the final substantial chapter was to explore a contrasting crime problem, which would present different modelling challenges, and this was done by considering the London riots of 2011. The topic is more coarse, in a number of senses, than those of the previous chapters: the relevant theory is specified in more general terms, the quality of data is relatively poor, and the patterns to be modelled are specified less precisely. The first component of the chapter was analytical, and revealed a number of general trends in the behaviour of offenders, from which three stages in the offending process were identified. These were used as the basis for a mathematical

model, which was constructed as a hybrid of three existing models. One primary consideration was the way in which aggregated effects - including feedback resulting from crowd behaviour - could be represented mathematically, and this informed the selection of the model. The model was tested using numerical simulation, and was found to be moderately successful in reproducing the observed characteristics of the London disorder. Using the parameter values established during that process, the model was then used to investigate a number of policy questions, with quantitative outcomes. Although intended only as a proof-of-concept, a more refined version of the model could be used for prediction or for scenario exploration.

7.2 Unifying themes

Although the issues arising within the topics studied were discussed at the end of each chapter, a number of more general themes emerge when considering the body of work as a whole. These include issues which occurred at multiple points within the thesis, together with observations relating to the more general concerns which arise in the analysis and modelling of systems such as these. In order to address these, a number of common topics will be identified and examined in this section.

7.2.1 Crime as a complex system

One of the foundations for the research presented in the thesis was the assertion that the processes which give rise to crime are, in general, complex, and require to be treated as such. Although precise definitions of complex systems vary, several principles are universal: they contain a large number of components which interact in a non-trivial way, giving rise to behaviour which cannot be understood fully by considering the components in isolation. Having explored a number of criminal issues, selected with this in mind, it is possible to assess whether their treatment as complex systems was justified.

The first chapter of original research, concerning event networks, is an example for which the above properties are demonstrated particularly clearly. In the event

network framework proposed, a link between two events not only indicates that they are close, but also implies that they are associated conceptually as common outcomes of a targeting process. These interactions are certainly non-trivial: criminological theory suggests that the occurrence of (near-)repeat incidents is a result of numerous factors, and so event network links are manifestations of a complex decision process. In addition, analysis of these networks demonstrated that the configuration of links is not regular and that certain signatures could be observed in their structure. This illustrates the value of considering the set of interactions as a whole, and indeed the presence of motifs and chains can be regarded as an emergent property. Such features arise from the accumulation of targeting decisions, but would not be seen if only pairwise relationships were considered. Pairwise analysis shows only that repeat incidents occur: using event networks, it is seen that ‘repeats of repeats’ have different character still, and that such sequences have lifetimes.

Characteristics of complexity are also evident in the modelling work of Chapters 5 and 6. In Chapter 5, it is shown that a simple diffusion model can give rise to ‘hot-spots’ of risk, which can also be considered to be an emergent property. Further evidence of this can be seen in the fact that the shape of these hot-spots can be reconciled with network structure: the manner in which risk is correlated across groups of streets could not be predicted *a priori*. The crowd effects in the riot model of Chapter 6 are also a classic example of emergence (and, indeed, this property has been investigated for several of the component models). The model encodes basic offender behaviours - participants make utility-maximising decisions based on simple calculations - yet gives rise to unified offending against a small number of targets.

Overall, the research presented has demonstrated that several crimes exhibit the characteristics of complex systems, in both their analysis and modelling. Furthermore, several meaningful findings are only made possible by the use of tools explicitly intended to incorporate these features. As well as justifying the general premise of the thesis, this suggests that such concerns should be accounted for in future research

into crime and its modelling.

7.2.2 The importance of realistic environmental backcloth

Principles of complexity are also evident in the analytical work concerning the street network in Chapters 3 and 4. The rationale for performing those analyses was that, for the purpose of crime analysis, it is necessary to consider the street network as a complex, integrated entity. According to theory, the influence of the street network on crime patterns is exerted through its role in shaping routine activities and, consequently, awareness spaces. In particular, the character of a particular street is hypothesised to be determined, at least in part, by its role in habitual travel through the network.

If the street network was not considered as a complex structure, individual streets could only be considered in terms of properties which were inherent to them, rather than in the context of the network as a whole. The deficiency of such an approach is demonstrated by the results of Chapters 3 and 4, in which a significant relationship was found between the incidence of crime and variables which account explicitly for the role of streets in the wider network. In this respect, the complexity lies in the structure of the city itself: its intricacy is the reason why offender awareness spaces are shaped non-trivially.

Importantly, these effects are evident for both the long-term spatial distribution of crime and for the phenomenon of (near-)repeat victimisation, which are the two principles upon which many predictive methods are based. There is good reason to conclude, therefore, that to fail to account for the street network in predictive models of crime represents a significant shortcoming. This is particularly acute when taking into account the emergent effects which arise when situating such a model on a network, as seen in Chapter 5.

The crucial issue in this context is that the features of interest arise only when

the network structure considered is realistic. Stylised or regular networks (such as grids) do not display these properties, and do not correspond to the real-world situations studied empirically. In order to be meaningful, therefore, the networks on which crime models are situated should reflect realistic structures. This accords with the more general trend towards the use of more realistic environmental representations in crime modelling, and indeed the network framework could be extended in line with this principle: network analyses could be adjusted to account for land use, for example.

7.2.3 Correspondence between theory and measurement

That the importance of street network effects can be claimed with confidence is partly due to the fact that the analyses which were carried out were designed to test criminological theory in the most direct way possible. This was done by selecting variables which correspond closely to the mechanisms by which the network is hypothesised to influence crime; *i.e.* those which estimate levels of awareness arising from habitual travel.

The majority of previous work concerning the street network has relied on proxy measures or those for which the correspondence with theory is opaque. This is partly for reasons of expedience, since more advanced measures require bespoke calculation. The gap between measurement and theory, however, means that the extent to which such analyses are a proper test of theory is open to question.

What has been demonstrated in this thesis is that it is possible to select, or design, metrics for which the relationship with hypothesised mechanisms can be delineated clearly. In the network context considered here, betweenness, which is well-known, and commonality, which was developed for this purpose, both measure the accumulation of journeys (albeit in a first-order way) and translate immediately to terms of pedestrian activity. In addition, both are quantitative and suitable for incorporation into mathematical models.

This example can be extended to a more general lesson for crime analysis. In criminological theory, concepts are frequently invoked for which no direct measure is immediately apparent, and tests of these involve some element of approximation or compromise. The case considered here demonstrates that it is possible to improve in this respect by developing novel measurements, and this may be applicable in other contexts. Analyses using such measures would be expected to carry greater weight as formal tests of theory.

7.2.4 Alternative modelling approaches

An issue which arises frequently in the modelling of crime, and social systems more generally, is the difficulty of representing theories which are specified at the level of the individual in mathematical terms. Both of the crimes for which models are presented here - residential burglary and rioting - draw on theories of that type, and the models were built from such a foundation. In the case of burglary, for example, it is assumed that aggregate patterns result from the accumulation of the activities of individual offenders, each of whom operates within an awareness space. The difficulty arises from the fact that awareness spaces are formed in the course of journeys around the urban area: there is no natural way to represent navigational heuristics in simple mathematical terms.

In cases such as these, a natural alternative is to adopt an agent-based approach, which allows for individual behaviours to be specified with unlimited complexity. Such models are popular within criminology for precisely this reason. One disadvantage of the approach, however, is that the analysis which can be performed on such a model is significantly less sophisticated than that which is possible for a traditional mathematical model. Given the complexity of the mechanisms to be modelled, though, it may appear in many cases that there is no alternative but to take an agent-based approach.

What is demonstrated by the two examples presented here, however, is that there are circumstances in which a continuous mathematical representation can be found for a system which appears initially to demand an agent-based approach. In the case of burglary, the key observation is that the agent-based aspect of the process - routine travel - is not dynamic at the same time-scale as the evolution of crime risk. When the system is framed as it is, the role of activity spaces is to influence the inherent potential for crime at a place; this is determined by the (static) structure of the network and does not vary in response to burglary events. This aspect of the model can therefore be calculated ‘offline’ (through the calculation of betweenness, which itself can be defined in agent-based terms) and the remainder of the model can be specified using differential equations.

The riot model was also specified, at the most basic level, in terms of the behaviour of individuals: all participants were modelled as making decisions based on their perception of the relative utility of various actions (*e.g.* target choices). The key step in the transition from an agent-based approach again lies in the basic structure of the model, in which the riot is described in aggregated terms: rather than tracking individual participants, the quantities which are modelled represent total stocks and flows. Individual behaviours are therefore averaged (using entropy-maximisation, for example) to produce simple equations for the aggregate quantities; essentially a mean-field approach. That this is a feasible approach is a result of a fact that, in contrast to the case of burglary, the aim is not to identify individual crimes, but to describe the overall character of the riot.

The overall conclusion here is that there are situations in which the fundamental tension between agent-based and classical mathematical approaches can be resolved. The transition between the two has been demonstrated in work elsewhere, typically in cases in which the agent-based formulation is simple; however, these examples demonstrate that careful construction of the model can also be used to circumvent the problem. It should be noted, however, that while this discussion assumes that

the ultimate goal is to avoid an agent-based approach, it is not the case that such models are inferior. Although this thesis approaches modelling from a mathematical perspective, agent-based approaches have clear value in exploratory theoretical work, for example.

7.2.5 Appropriate null models

One final point which recurred within the thesis was the use, and selection, of appropriate null models during statistical analysis. This is primarily a technical point, but has relevance for the study of criminal phenomena more generally. The main observation is that, when analysing some novel property of crime data, of equal importance (and possibly difficulty) as the calculation of the property itself is the determination of a reference distribution which truly reflects the absence of the phenomenon under investigation.

This principle is best demonstrated by example, and was evident in Chapters 2, on event networks, and 4, on (near-)repeat target choice. In the analysis of event networks, the construction of the networks themselves and the counting of their features (*e.g.* motifs and chains) was relatively simple; the primary technical challenge was in finding distributions against which these frequencies could be compared. In doing so, it was necessary to produce networks with certain prescribed features while ensuring that they were valid representations of spatio-temporal data. The process used was particularly involved, but the value of the technique depends on it.

The issue of the null model was also evident in the work on commonality and (near-)repeats. When analysing the choices of offenders, it was essential to account for the opportunities which were *not* taken, since not all sets of choices were equivalent. This is partly due to variation in commonality: a given value might be among the highest in one choice set, but among the lowest in another. A discrete choice model was chosen for this reason, and made it possible to account for the relative nature of these choices, but this illustrates the general principle.

The main observation from this discussion concerns the importance of null models in the development of novel analyses: event networks and commonality are of little value if their significance cannot be evaluated.

7.3 Further work

As well as the avenues for further work which were identified at the end of each chapter, a number of opportunities concern either the unification of aspects of the work, or open technical questions which are relevant to several aspects. These will be identified in this section.

7.3.1 Implementation and integration of predictive systems

Much of the work presented in the thesis is either concerned explicitly with crime prediction or has the potential to inform it. Although a number of positive results have been found, in the sense that useful relationships have been shown to exist, the question of how these relate to each other, and how they might be combined, remains open. This is an issue which must be resolved if such methods are to be deployed in the real world.

This issue is most pertinent for the work concerned with the crime of residential burglary, since it incorporates a number of complementary approaches. Importantly, many of the principles involved, such as the influence of urban form on crime, are general, so that insights relating to prediction might be applicable more widely.

The most immediate example of a predictive tool is the model developed in Chapter 5, and this builds directly on the empirical findings of Chapters 3 and 4. Regardless of its validity or properties, however, a number of issues are required to be resolved before it could be deployed operationally. Among these are the related questions of how the model is calibrated and how its outputs are to be used in informing real-world interventions.

The objective of calibration is, of course, to optimise the behaviour of the model so that it describes real-world behaviour as closely as possible. The definition of ‘optimal’, however, depends on how results are to be used: since the objective is prediction (and, beyond that, the prevention of crime) the results should be evaluated in those terms. A natural way to do this is to optimise the number of crimes occurring on the streets with the highest modelled risk; however, there is still uncertainty in how this should be done.

Since many of the intended interventions (*e.g.* foot patrol) will be resource-limited, it is logical that the optimisation of predictive accuracy should be performed with respect to what is possible operationally. There is little value in optimising the number of crimes occurring in the top 100 street segments, for example, if only 10 can feasibly be patrolled. The situation is complicated further still by the configuration of the network: a contiguous set of streets can be patrolled much more easily than a set which are well-spread. Optimisation under these circumstances is the next challenge in the development of the model.

A further question concerns how the work on event networks, which also has potential as a predictive tool, could be incorporated in this framework. The strength of that approach lies in its nuanced treatment of patterns of events, and this could be used to refine the diffusion model. One possibility would be to allow the ‘boost’ in risk resulting from an offence to vary according to whether the offence was part of a longer chain. Since it has been shown that the probability of a follow-up offence varies according to how many offences have preceded it, there is good reason to incorporate this. A similar idea arises from the analysis of motifs: if two close events have occurred, risk could be forced to diffuse towards their common neighbourhood (as implied by the most prevalent motif). In this way, the two approaches could be combined to produce an integrated predictive system.

7.3.2 Parameter selection

A further issue which arose on a number of occasions in the thesis is also related to calibration, and concerns the appropriateness of using point estimates of parameters. The difficulties occur for one of two reasons: either because a range of values can all be regarded as valid, or because it is not possible to discriminate between the results arising from different settings. Examples of each of these were encountered during the thesis.

In Chapter 3, the network metric betweenness was calculated on a radially-limited basis; that is, an upper limit was placed on the length of path which contributed to the calculation. This length therefore acts as a parameter for betweenness, and a range of values were considered in the statistical analysis. Since statistical significance was evident for almost all values considered, the question arises as to which version of betweenness is most meaningful (for predictive purposes or otherwise). In this context, selecting one particular value (that for which the z -score is highest, for example) could be considered to be reductive and arbitrary. Although each value is a variation on the same basic concept, there is little reason to say that one is definitive. In fact, it may be the case that different versions of betweenness predict different offences: longer radii may correspond to more exploratory offending, for example.

A similar issue was also encountered in the riot modelling of Chapter 6. In that case, model outputs were assessed to be acceptable if they satisfied a number of criteria. These criteria were sufficiently broad (for reasons articulated in that chapter) that a range of parameter configurations gave rise to acceptable outputs; again, therefore, the question arose of how to select from this. Although a point estimate was used for the purpose of the analysis which followed, the lack of a precise desired output means there is little reason to favour one configuration over another.

It is clear that resolving these issues is an important step in moving forward with

this work, and there are several ways in which this could be done. One possibility is to use weighted combinations of values: in the case of betweenness, this would mean taking the average value across a number of different radii. Similarly in principle, a stochastic approach could be used: whenever required, the parameter in question would be randomly sampled from the acceptable range.

An alternative approach would be to produce results (*i.e.* statistical estimates or model outputs) for all acceptable parameter values, and use these to calculate confidence intervals for the point-estimated results. This is similar in spirit to sensitivity analysis, but confronts a subtly different problem. Assessing the relative strength of these approaches is an important next step in the development of these analyses and models.

7.3.3 Effect of policing

One consideration which is conspicuously absent from much of the modelling work presented here is that of the effect of policing. In the case of burglary, policing is absent from the model, and the evolution of risk is due to either the occurrence of offences or the natural decay over time. In the riot model, policing is present - such activity is prominent in the decisions of participants - but its representation is over-simplistic. The reason for both of these is a more general lack of understanding of the behaviour of police and its effect on crime.

In comparison with other phenomena, the behaviour and effect of policing is not well-understood within criminology. In terms of police activity, there is little empirical work concerning the coverage of routine patrols, their frequency, and how they are determined; in any case, such results would be unlikely to be universal. This is of particular relevance given the focus of this work on the street network, since it is unknown whether more central streets are favoured by police patrols, for example. The same is true for rioting: there is a lack of data concerning the police response to such incidents.

A further open problem, related to this, concerns the effect of policing interventions on crime. It is generally unknown, for example, whether movement of police through an area has an effect in reducing crime and, if so, what magnitude and persistence this effect has. Because of this, there is little basis on which to incorporate the effect in a model. The same is true for rioting: without more accurate data, it is impossible to estimate the deterrent effect of policing.

The only way in which this problem can be addressed is through empirical research, which, in turn, requires increased availability of data on the activity of police. Fortunately, this availability is increasing, as the use of GPS tracking devices increases within police forces. With access to this data, it should be possible to gain insight into police movements and, by comparing these with the occurrence of crime, quantify the extent to which crime is deterred. It is at this stage that such effect can be incorporated into models in a well-motivated way, so that these models can then be used to guide the allocation of resources in a real-world setting.

References

- Almendral, J. A., & Díaz-Guilera, A. (2007). Dynamical and spectral properties of complex networks. *New Journal of Physics*, 9(6), 187.
- Anderson, P. W. (1972). More is different. *Science*, 177(4047), 393–396.
- Anderson, R. M., & May, R. M. (1992). *Infectious diseases of humans: Dynamics and control*. Oxford, UK: Oxford University Press.
- Andresen, M. A. (2014). *Environmental criminology: Evolution, theory, and practice*. Abingdon, UK: Routledge.
- Anthonisse, J. M. (1971). The rush in a directed graph. Technical Report BN 9/71, Stichting Mathematisch Centrum, Amsterdam.
- Arenas, A., Díaz-Guilera, A., & Pérez-Vicente, C. J. (2006a). Synchronization processes in complex networks. *Physica D: Nonlinear Phenomena*, 224(1–2), 27–34.
- Arenas, A., Díaz-Guilera, A., & Pérez-Vicente, C. J. (2006b). Synchronization reveals topological scales in complex networks. *Physical Review Letters*, 96(11), 114102.
- Armitage, R. (2007). Sustainability versus safety: Confusion, conflict and contradiction in designing out crime. In G. Farrell, K. J. Bowers, S. D. Johnson, & M. Townsley (eds.) *Imagination for crime prevention: Essays in honour of Ken Pease*, vol. 21 of *Crime Prevention Studies*, pp. 81–110. Monsey, NY: Criminal Justice Press.
- Arrowsmith, D., & Place, C. (1992). *Dynamical systems: Differential equations, maps, and chaotic behaviour*. Boca Raton, FL: Chapman & Hall/CRC.

- Ashby, M. P., & Bowers, K. J. (2013). A comparison of methods for temporal analysis of aoristic crime. *Crime Science*, 2(1), 1.
- Ashton, J., Brown, I., Senior, B., & Pease, K. (1998). Repeat victimisation: Offender accounts. *International Journal of Risk Security and Crime Prevention*, 3(4), 269–280.
- Auyero, J., & Moran, T. P. (2007). The dynamics of collective violence: Dissecting food riots in contemporary Argentina. *Social Forces*, 85(3), 1341–1367.
- Ballester, C., Calvó-Armengol, A., & Zenou, Y. (2006). Who’s who in networks. Wanted: The key player. *Econometrica*, 74(5), 1403–1417.
- Bang-Jensen, J., & Gutin, G. (2007). *Digraphs: Theory, algorithms and applications*. Springer Monographs in Mathematics. London, UK: Springer.
- Barahona, M., & Pecora, L. M. (2002). Synchronization in small-world systems. *Physical Review Letters*, 89(5), 054101.
- Barr, R., & Pease, K. (1990). Crime placement, displacement, and deflection. *Crime and Justice*, 12, 277–318.
- Barthélemy, M. (2011). Spatial networks. *Physics Reports*, 499(1–3), 1–101.
- Barthélemy, M., & Flammini, A. (2008). Modeling urban street patterns. *Physical Review Letters*, 100(13), 138702.
- Barthélemy, M., Nadal, J.-P., & Berestycki, H. (2010). Disentangling collective trends from local dynamics. *Proceedings of the National Academy of Sciences*, 107(17), 7629–7634.
- Batty, M. (2007). *Cities and complexity: Understanding cities with cellular automata, agent-based models, and fractals*. Cambridge, MA: MIT Press.
- Baudains, P., Braithwaite, A., & Johnson, S. D. (2013a). Spatial patterns in the 2011 London riots. *Policing*, 7(1), 21–31.

- Baudains, P., Braithwaite, A., & Johnson, S. D. (2013b). Target choice during extreme events: A discrete spatial choice model of the 2011 London riots. *Criminology*, *51*(2), 251–285.
- Beavon, D. J. K., Brantingham, P. L., & Brantingham, P. J. (1994). The influence of street networks on the patterning of property offences. In R. V. Clarke (ed.) *Crime prevention studies*, vol. 2, pp. 115–148. Monsey, NY: Criminal Justice Press.
- Bellomo, N., Berestycki, H., Brezzi, F., & Nadal, J.-P. (2009). Mathematics and complexity in life and human sciences. *Mathematical Models and Methods in Applied Sciences*, *20*(Supplement), 1385–1389.
- Ben-Galim, D., & Gottfried, G. (2011). Exploring the relationship between riot areas and deprivation: An IPPR analysis. London, UK: Institute for Public Policy Research.
- Berestycki, H., Johnson, S. D., Ockendon, R., & Primicerio, M. (2010). Foreword to the special issue of EJAM on crime modelling. *European Journal of Applied Mathematics*, *21*(Special Double Issue 4-5), 271–274.
- Berestycki, H., & Nadal, J.-P. (2010). Self-organised critical hot spots of criminal activity. *European Journal of Applied Mathematics*, *21*(Special Double Issue 4-5), 371–399.
- Berestycki, H., Rodríguez, N., & Ryzhik, L. (2013). Traveling wave solutions in a reaction-diffusion model for criminal activity. *Multiscale Modeling & Simulation*, *11*(4), 1097–1126.
- Berk, R., & Macdonald, J. (2009). The dynamics of crime regimes. *Criminology*, *47*(3), 971–1008.
- Berk, R. A. (1974). A gaming approach to crowd behavior. *American Sociological Review*, *39*(3), 355–373.
- Berlekamp, E. R., Conway, J. H., & Guy, R. K. (2003). *Winning ways for your mathematical plays*. Natick, MA: A K Peters.

- Bernasco, W. (2008). Them again? Same-offender involvement in repeat and near repeat burglaries. *European Journal of Criminology*, 5(4), 411–431.
- Bernasco, W. (2009). Modeling micro-level crime location choice: Application of the discrete choice framework to crime at places. *Journal of Quantitative Criminology*, 26(1), 113–138.
- Bernasco, W., & Block, R. (2009). Where offenders choose to attack: A discrete choice model of robberies in Chicago. *Criminology*, 47(1), 93–130.
- Bernasco, W., & Block, R. (2011). Robberies in Chicago: A block-level analysis of the influence of crime generators, crime attractors, and offender anchor points. *Journal of Research in Crime and Delinquency*, 48(1), 33–57.
- Bernasco, W., & Nieuwbeerta, P. (2005). How do residential burglars select target areas? *British Journal of Criminology*, 45(3), 296–315.
- Bevis, C., & Nutter, J. (1977). *Changing street layouts to reduce residential burglary*. Minneapolis, MN: Minnesota Crime Prevention Center.
- Bhat, S. S., & Maciejewski, A. A. (2006). An agent-based simulation of the LA 1992 riots. In *Conference on artificial intelligence| ICAI*, vol. 6, p. 76.
- Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. In *Proceedings of the 3rd Swiss Transport Research Conference*. Ascona, Switzerland.
- Birks, D., Townsley, M., & Stewart, A. (2012). Generative explanations of crime: Using simulation to test criminological theory. *Criminology*, 50(1), 221–254.
- Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*. New York, NY: Free Press.
- Block, C., Dabdou, M., & Fregly, S. (1995). *Crime analysis through computer mapping*. Washington, DC: Police Executive Research Forum.

- Block, R. L., & Block, C. R. (1995). Space, place, and crime: Hot spot areas and hot places of liquor-related crime. In J. E. Eck, & D. Weisburd (eds.) *Crime and place*, vol. 4 of *Crime Prevention Studies*, pp. 145–183. Monsey, NY: Criminal Justice Press.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424(4–5), 175–308.
- Bollobás, B. (2002). *Modern graph theory*. New York, NY: Springer, 2nd ed.
- Bowers, K. J., & Johnson, S. D. (2004). Who commits near repeats? A test of the boost explanation. *Western Criminology Review*, 5(3), 12–24.
- Bowers, K. J., & Johnson, S. D. (2005). Domestic burglary repeats and space-time clusters. *European Journal of Criminology*, 2(1), 67–92.
- Bowers, K. J., Johnson, S. D., & Pease, K. (2004). Prospective hot-spotting: The future of crime mapping? *British Journal of Criminology*, 44(5), 641–658.
- Braga, A. A. (2001). The effects of hot spots policing on crime. *The Annals of the American Academy of Political and Social Science*, 578(1), 104–125.
- Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Social Networks*, 30(2), 136–145.
- Brantingham, P. J., & Brantingham, P. L. (1981). *Environmental criminology*. Beverly Hills, CA: Sage.
- Brantingham, P. J., & Tita, G. (2008). Offender mobility and crime pattern formation from first principles. In L. Liu, & J. Eck (eds.) *Artificial crime analysis systems: Using computer simulations and geographic information systems*, pp. 193–208. Hershey, PA: IGI Global.
- Brantingham, P. L., & Brantingham, P. J. (1975). Residential burglary and urban form. *Urban Studies*, 12(3), 273–284.

- Brantingham, P. L., & Brantingham, P. J. (1993a). Environment, routine and situation: Toward a pattern theory of crime. *Advances in Criminological Theory*, 5, 259–294.
- Brantingham, P. L., & Brantingham, P. J. (1993b). Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology*, 13(1), 3–28.
- Brantingham, P. L., & Brantingham, P. J. (1995). Criminality of place. *European Journal on Criminal Policy and Research*, 3(3), 5–26.
- Brantingham, P. L., Kinney, B., Glässer, U., Jackson, P., & Vajihollahi, M. (2008). Mastermind: Computational modeling and simulation of spatiotemporal aspects of crime in urban environments. In L. Liu, & J. Eck (eds.) *Artificial crime analysis systems: Using computer simulations and geographic information systems*, pp. 252–280. Hershey, PA: IGI Global.
- Breu, H., & Kirkpatrick, D. G. (1998). Unit disk graph recognition is NP-hard. *Computational Geometry*, 9(1–2), 3–24.
- Brunsdon, C., Corcoran, J., & Higgs, G. (2007). Visualising space and time in crime patterns: A comparison of methods. *Computers, Environment and Urban Systems*, 31(1), 52–75.
- Budd, T. (1999). *Burglary of domestic dwellings: Findings from the British Crime Survey*. No. 4/99 in Home Office Statistical Bulletin. London, UK: Home Office.
- Buhl, J., Gautrais, J., Reeves, N., Solé, R. V., Valverde, S., Kuntz, P., & Theraulaz, G. (2006). Topological patterns in street networks of self-organized urban settlements. *The European Physical Journal B - Condensed Matter and Complex Systems*, 49(4), 513–522.
- Cantrell, R. S., Cosner, C., & Manásevich, R. (2012). Global bifurcation of solutions for crime modeling equations. *SIAM Journal on Mathematical Analysis*, 44(3), 1340–1358.

- Cardillo, A., Scellato, S., Latora, V., & Porta, S. (2006). Structural properties of planar graphs of urban street patterns. *Physical Review E*, 73(6), 066107.
- Chainey, S. P., & Ratcliffe, J. H. (2005). *GIS and crime mapping*. London, UK: Wiley.
- Chan, S. H. Y., Donner, R. V., & Lämmer, S. (2011). Urban road networks — spatial networks with universal geometric features? *The European Physical Journal B*, 84(4), 563–577.
- Cheng, X.-Q., & Shen, H.-W. (2010). Uncovering the community structure associated with the diffusion dynamics on networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(4), P04024.
- Chung, F. R. (1997). *Spectral graph theory*. No. 92 in CBMS Regional Conference Series in Mathematics. Providence, RI: AMS.
- Clare, J., Fernandez, J., & Morgan, F. (2009). Formal evaluation of the impact of barriers and connectors on residential burglars' macro-level offending location choices. *Australian & New Zealand Journal of Criminology*, 42(2), 139–158.
- Clark, B. N., Colbourn, C. J., & Johnson, D. S. (1990). Unit disk graphs. *Discrete Mathematics*, 86(1–3), 165–177.
- Clarke, R. V., & Weisburd, D. (1994). Diffusion of crime control benefits: Observations on the reverse of displacement. In R. V. Clarke (ed.) *Crime prevention studies*, vol. 2, pp. 165–183. Monsey, NY: Criminal Justice Press.
- Clarke, R. V. G. (1980). Situational crime prevention: Theory and practice. *British Journal of Criminology*, 20(2), 136–147.
- Clarke, R. V. G. (2004). *Closing streets and alleys to reduce crime: Should you go down this road?*. Washington, DC: US Department of Justice, Office of Community Oriented Policing Services.

- Cohen, J., & Tita, G. (1999). Diffusion in homicide: Exploring a general method for detecting spatial diffusion processes. *Journal of Quantitative Criminology*, 15(4), 451–493.
- Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44(4), 588–608.
- Conroy Dalton, R. (2003). The secret is to follow your nose: Route path selection and angularity. *Environment and Behavior*, 35(1), 107–131.
- Cormen, T. H. (2009). *Introduction to algorithms*. Cambridge, MA: MIT Press.
- Cornish, D. B., & Clarke, R. V. G. (1986). *The reasoning criminal: Rational choice perspectives on offending*. New York, NY: Springer Verlag.
- Costa, L. d. F., Oliveira, O. N., Travieso, G., Rodrigues, F. A., Villas Boas, P. R., Antiqueira, L., Viana, M. P., & Correa Rocha, L. E. (2011). Analyzing and modeling real-world phenomena with complex networks: A survey of applications. *Advances in Physics*, 60(3), 329–412.
- Cromwell, P. F., Olson, J. N., & Avary, D. W. (1991). *Breaking and entering: An ethnographic analysis of burglary*, vol. 8 of *Studies in Crime, Law and Justice*. Newbury Park, CA: Sage.
- Crucitti, P., Latora, V., & Porta, S. (2006a). Centrality in networks of urban streets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(1), 015113.
- Crucitti, P., Latora, V., & Porta, S. (2006b). Centrality measures in spatial networks of urban streets. *Physical Review E*, 73(3), 036125.
- Department for Communities and Local Government (2010). *English indices of deprivation 2010: Guidance*. London, UK: Department for Communities and Local Government.
- Department for Communities and Local Government (2011). *Extent of retail development taking place within town centres in England*. London, UK: Department for Communities and Local Government.

- Diestel, R. (2010). *Graph theory*. No. 173 in Graduate Texts in Mathematics. Heidelberg, Germany: Springer-Verlag, 4th ed.
- Eck, J. E. (2002). Learning from experience in problem-oriented policing and situational prevention: The positive functions of weak evaluations and the negative functions of strong ones. In N. Tilley (ed.) *Evaluation for crime prevention*, vol. 14 of *Crime Prevention Studies*, pp. 93–118. Monsey, NY: Criminal Justice Press.
- Elffers, H., & Van Baal, P. (2008). Realistic spatial backcloth is not that important in agent based simulation research: An illustration from simulating perceptual deterrence. In L. Liu, & J. Eck (eds.) *Artificial crime analysis systems: Using computer simulations and geographic information systems*, pp. 19–34. Hershey, PA: IGI Global.
- Epstein, J. M. (2002). Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences*, 99(Supplement 3), 7243–7250.
- Epstein, J. M. (2006). *Generative social science: Studies in agent-based computational modeling*. Princeton, NJ: Princeton University Press.
- Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up*. Washington, DC: Brookings Institution Press.
- Euler, L. (1741). Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, 8, 128–140.
- Everson, S., & Pease, K. (2001). Crime against the same person and place: Detection opportunity and offender targeting. In G. Farrell, & K. Pease (eds.) *Repeat victimization*, vol. 12 of *Crime Prevention Studies*, pp. 199–220. Monsey, NY: Criminal Justice Press.
- Farrell, G. (2005). Progress and prospects in the prevention of repeat victimization. In N. Tilley (ed.) *Handbook of crime prevention and community safety*, pp. 145–172. Cullompton, UK: Willan.

- Farrell, G., & Pease, K. (1994). Crime seasonality: Domestic disputes and residential burglary in Merseyside 1988–90. *British Journal of Criminology*, 34(4), 487–498.
- Farrell, G., Phillips, C., & Pease, K. (1995). Like taking candy: Why does repeat victimization occur? *British Journal of Criminology*, 35(3), 384–399.
- Farrington, D. P. (1983). Offending from 10 to 25 years of age. In K. T. Van Dusen, & S. A. Mednick (eds.) *Prospective studies of crime and delinquency*, no. 2 in Longitudinal Research in the Behavioral, Social and Medical Sciences, pp. 17–37. Dordrecht, Netherlands: Kluwer-Nijhoff.
- Felson, M. (1987). Routine activities and crime prevention in the developing metropolis. *Criminology*, 25(4), 911–932.
- Felson, M., & Poulsen, E. (2003). Simple indicators of crime by time of day. *International Journal of Forecasting*, 19(4), 595–601.
- Field, S. (1992). The effect of temperature on crime. *British Journal of Criminology*, 32(3), 340–351.
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3–5), 75–174.
- Frank, R., Andresen, M. A., & Brantingham, P. L. (2012). Criminal directionality and the structure of urban form. *Journal of Environmental Psychology*, 32(1), 37–42.
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35–41.
- Freeman, L. C. (1978). Centrality in social networks: Conceptual clarification. *Social Networks*, 1(3), 215–239.
- Freud, S. (1957). *Group psychology and the analysis of ego*. London, UK: International Psychoanalytic Press.
- Garoupa, N. (1997). The theory of optimal law enforcement. *Journal of Economic Surveys*, 11(3), 267–295.

- Gilbert, N. (2008). *Agent-based models*, vol. 153 of *Quantitative Applications in the Social Sciences*. London, UK: Sage Publications.
- Gini, C. (1921). Measurement of inequality of incomes. *The Economic Journal*, 31(121), 124–126.
- Goh, C. K., Quek, H. Y., Tan, K. C., & Abbass, H. A. (2006). Modeling civil violence: An evolutionary multi-agent, game theoretic approach. In D. B. Fogel (ed.) *Proceedings of the 2006 IEEE Congress on Evolutionary Computation*, pp. 1624–1631. Vancouver, Canada.
- Gordon, M. B. (2010). A random walk in the literature on criminality: A partial and critical view on some statistical analyses and modelling approaches. *European Journal of Applied Mathematics*, 21(Special Double Issue 4-5), 283–306.
- Gordon, M. B., Iglesias, J. R., Semeshenko, V., & Nadal, J.-P. (2009). Crime and punishment: The economic burden of impunity. *The European Physical Journal B*, 68(1), 133–144.
- Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83(6), 1420–1443.
- Groff, E. R. (2007a). Simulation for theory testing and experimentation: An example using routine activity theory and street robbery. *Journal of Quantitative Criminology*, 23(2), 75–103.
- Groff, E. R. (2007b). ‘Situating’ simulation to model human spatio-temporal interactions: An example using crime events. *Transactions in GIS*, 11(4), 507–530.
- Groff, E. R., & La Vigne, N. G. (2002). Forecasting the future of predictive crime mapping. In N. Tilley (ed.) *Analysis for crime prevention*, vol. 13 of *Crime Prevention Studies*, pp. 29–57. Cullompton, UK: Willan.
- Gross, M. (2011). Why do people riot? *Current Biology*, 21(18), 673–676.
- Grubestic, T. H., & Mack, E. A. (2008). Spatio-temporal interaction of urban crime. *Journal of Quantitative Criminology*, 24(3), 285–306.

- Gunderson, L., & Brown, D. (2000). Using a multi-agent model to predict both physical and cyber criminal activity. In *Proceedings of the 2000 IEEE International Conference on Systems, Man, and Cybernetics*, vol. 4, pp. 2338–2343.
- Haldane, A. G., & May, R. M. (2011). Systemic risk in banking ecosystems. *Nature*, 469(7330), 351–355.
- Harris, B., & Wilson, A. G. (1978). Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models. *Environment and Planning A*, 10(4), 371–388.
- Hawley, A. H. (1950). *Human ecology: A theory of community structure*. New York, NY: Ronald.
- Hayslett-McCall, K. L., Qui, F., Curtin, K. M., Chastain, B., Schubert, J., & Carver, V. (2008). The simulation of the journey to residential burglary. In L. Liu, & J. Eck (eds.) *Artificial crime analysis systems: Using computer simulations and geographic information systems*, pp. 281–300. Hershey, PA: IGI Global.
- Hillier, B. (1996). *Space is the machine*. Cambridge, UK: Cambridge University Press.
- Hillier, B. (2004). Can streets be made safe? *Urban Design International*, 9(1), 31–45.
- Hillier, B., & Hanson, J. (1984). *The social logic of space*. Cambridge, UK: Cambridge University Press.
- Hillier, B., & Shu, S. (2000). Crime and urban layout: The need for evidence. In S. Ballantyne, K. Pease, & V. MacLaren (eds.) *Secure foundations: Key issues in crime prevention, crime reduction and community safety*, pp. 224–248. London, UK: Institute for Public Policy Research.
- Hipp, J. R., Curran, P. J., Bollen, K. A., & Bauer, D. J. (2004). Crimes of opportunity or crimes of emotion? Testing two explanations of seasonal change in crime. *Social Forces*, 82(4), 1333–1372.

- Hirschfield, A., & Bowers, K. J. (1997). The effect of social cohesion on levels of recorded crime in disadvantaged areas. *Urban Studies*, 34(8), 1275–1295.
- HMIC (2011). *The rules of engagement: A review of the August 2011 disorders*. London, UK: Her Majesty's Inspectorate of Constabulary.
- Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519(3), 97–125.
- Home Office (2011). *An overview of recorded crimes and arrests resulting from disorder events in August 2011*. London, UK: Home Office.
- Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1), 253–258.
- Iwanski, N., Frank, R., Dabbaghian, V., Reid, A., & Brantingham, P. (2011). Analyzing an offender's journey to crime: A criminal movement model (CriMM). In *Proceedings of the 2011 European Intelligence and Security Informatics Conference*, pp. 70–77.
- Iwanski, N., Frank, R., Reid, A., & Dabbaghian, V. (2012). A computational model for predicting the location of crime attractors on a road network. In *Proceedings of the 2012 European Intelligence and Security Informatics Conference*, pp. 60–67.
- Jacobs, J. (1961). *The life and death of great American cities*. New York, NY: Random House.
- Jacquez, G. M. (1996). A k nearest neighbour test for space-time interaction. *Statistics in Medicine*, 15(18), 1935–1949.
- Jeffery, C. R. (1971). *Crime prevention through environmental design*. Beverly Hills, CA: Sage.
- Jiang, B. (2007). A topological pattern of urban street networks: Universality and peculiarity. *Physica A: Statistical Mechanics and its Applications*, 384(2), 647–655.

- Jiang, B., & Claramunt, C. (2004). Topological analysis of urban street networks. *Environment and Planning B: Planning and Design*, 31(1), 151 – 162.
- Johnson, N. F. (2007). *Two's company, three is complexity: A simple guide to the science of all sciences*. Oxford, UK: Oneworld.
- Johnson, N. F., Jefferies, P., & Hui, P. M. (2003). *Financial market complexity*. Oxford, UK: Oxford University Press.
- Johnson, S. D. (2008). Repeat burglary victimisation: A tale of two theories. *Journal of Experimental Criminology*, 4(3), 215–240.
- Johnson, S. D. (2014). How do offenders choose where to offend? Perspectives from animal foraging. *Legal and Criminological Psychology*, 19(2), 193–210.
- Johnson, S. D., Bernasco, W., Bowers, K. J., Elffers, H., Ratcliffe, J., Rengert, G., & Townsley, M. (2007). Space–time patterns of risk: A cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology*, 23(3), 201–219.
- Johnson, S. D., Bowers, K., & Hirschfield, A. (1997). New insights into the spatial and temporal distribution of repeat victimization. *British Journal of Criminology*, 37(2), 224–241.
- Johnson, S. D., & Bowers, K. J. (2004a). The burglary as clue to the future: The beginnings of prospective hot-spotting. *European Journal of Criminology*, 1(2), 237–255.
- Johnson, S. D., & Bowers, K. J. (2004b). The stability of space-time clusters of burglary. *British Journal of Criminology*, 44(1), 55–65.
- Johnson, S. D., & Bowers, K. J. (2007). Burglary prediction: Theory, flow and friction. In G. Farrell, K. J. Bowers, S. D. Johnson, & M. Townsley (eds.) *Imagination for crime prevention: Essays in honour of Ken Pease*, vol. 21 of *Crime Prevention Studies*, pp. 203–223. Monsey, NY: Criminal Justice Press.

- Johnson, S. D., & Bowers, K. J. (2010). Permeability and burglary risk: Are cul-de-sacs safer? *Journal of Quantitative Criminology*, 26(1), 89–111.
- Johnson, S. D., Bowers, K. J., Birks, D. J., & Pease, K. (2009a). Predictive mapping of crime by ProMap: Accuracy, units of analysis, and the environmental backcloth. In D. Weisburd, W. Bernasco, & G. J. Bruinsma (eds.) *Putting crime in its place*, pp. 171–198. New York, NY: Springer.
- Johnson, S. D., & Braithwaite, A. (2009). Spatio-temporal modelling of insurgency in Iraq. In J. D. Freilich, & G. R. Newman (eds.) *Reducing terrorism through situational crime prevention*, pp. 9–32. New York, NY: Criminal Justice Press.
- Johnson, S. D., Lab, S. P., & Bowers, K. J. (2008). Stable and fluid hotspots of crime: Differentiation and identification. *Built Environment*, 34(1), 32–45.
- Johnson, S. D., Summers, L., & Pease, K. (2009b). Offender as forager? A direct test of the boost account of victimization. *Journal of Quantitative Criminology*, 25, 181–200.
- Jones, P. A., Brantingham, P. J., & Chayes, L. R. (2010). Statistical models of criminal behavior: The effects of law enforcement actions. *Mathematical Models and Methods in Applied Sciences*, 20(Supplement), 1397–1423.
- Kalapala, V., Sanwalani, V., Clauset, A., & Moore, C. (2006). Scale invariance in road networks. *Physical Review E*, 73(2), 026130.
- Kashtan, N., Itzkovitz, S., Milo, R., & Alon, U. (2004). Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics*, 20(11), 1746–1758.
- Kirkland, J. A., & Maciejewski, A. A. (2003). A simulation of attempts to influence crowd dynamics. In *Proceedings of the 2003 IEEE International Conference on Systems, Man and Cybernetics*, vol. 5, pp. 4328 – 4333.
- Knox, G. (1964). Epidemiology of childhood leukaemia in Northumberland and Durham. *British Journal of Preventive and Social Medicine*, 18(1), 17–24.

- Knutsson, J. (2009). Standards of evaluations in problem-oriented policing projects: Good enough? In J. Knutsson, & N. Tilley (eds.) *Evaluating crime reduction initiatives*, vol. 24 of *Crime Prevention Studies*, pp. 7–28. Cullompton, UK: Willan.
- Kolokolnikov, T., Ward, M., & Wei, J. (2014). The stability of steady-state hot-spot patterns for a reaction-diffusion model of urban crime. *Discrete and Continuous Dynamical Systems - Series B*, 19(5), 1373–1410.
- Lämmer, S., Gehlsen, B., & Helbing, D. (2006). Scaling laws in the spatial structure of urban road networks. *Physica A: Statistical Mechanics and its Applications*, 363(1), 89–95.
- Laycock, G. (2005). Defining crime science. In M. J. Smith, & N. Tilley (eds.) *Crime science: New approaches to preventing and detecting crime*, pp. 3–24. Cullompton, UK: Willan.
- Le Bon, G. (1960). *The crowd: A study of the popular mind*. New York, NY: Viking Press.
- Lersch, K. M. (2004). *Space, time, and crime*. Durham, NC: Carolina Academic Press.
- Levin, S. A. (1998). Ecosystems and the biosphere as complex adaptive systems. *Ecosystems*, 1(5), 431–436.
- Lim, M., Metzler, R., & Bar-Yam, Y. (2007). Global pattern formation and ethnic/cultural violence. *Science*, 317(5844), 1540–1544.
- Liu, H., & Brown, D. E. (2003). Criminal incident prediction using a point-pattern-based density model. *International Journal of Forecasting*, 19(4), 603–622.
- Liu, L., & Eck, J. (2008). *Artificial crime analysis systems: Using computer simulations and geographic information systems*. Hershey, PA: IGI Global.
- Liu, L., Wang, X., Eck, J., & Liang, J. (2005). Simulating crime events and crime patterns in RA/CA model. In F. Wang (ed.) *Geographic information systems and crime analysis*, pp. 197–213. Hershey, PA: Idea Group Publishing.

- Lloyd, D. J., & O'Farrell, H. (2013). On localised hotspots of an urban crime model. *Physica D: Nonlinear Phenomena*, 253, 23–39.
- Malleson, N., Heppenstall, A., & See, L. (2010). Crime reduction through simulation: An agent-based model of burglary. *Computers, Environment and Urban Systems*, 34, 236–250.
- Malleson, N., See, L., Evans, A., & Heppenstall, A. (2012). Implementing comprehensive offender behaviour in a realistic agent-based model of burglary. *Simulation*, 88(1), 50–71.
- Mangan, S., & Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences*, 100(21), 11980–11985.
- Mantegna, R. N., & Stanley, H. E. (2000). *An introduction to econophysics: Correlations and complexity in finance*. Cambridge, UK: Cambridge University Press.
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2), 209–220.
- Marathe, M. V., Breu, H., Hunt, H. B., Ravi, S. S., & Rosenkrantz, D. J. (1995). Simple heuristics for unit disk graphs. *Networks*, 25(2), 59–68.
- Marchione, E., & Johnson, S. D. (2013). Spatial, temporal and spatio-temporal patterns of maritime piracy. *Journal of Research in Crime and Delinquency*, 50(4), 504–524.
- Mason, T. D. (1984). Individual participation in collective racial violence: A rational choice synthesis. *The American Political Science Review*, 78(4), 1040–1056.
- Masucci, A. P., Smith, D., Crooks, A., & Batty, M. (2009). Random planar graphs and the London street network. *The European Physical Journal B*, 71(2), 259–271.
- Mawby, R. I. (1977). Defensible space: A theoretical and empirical appraisal. *Urban Studies*, 14(2), 169–179.

- May, R. M. (1972). Will a large complex system be stable? *Nature*, 238(5364), 413–414.
- May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261(5560), 459–467.
- McCord, E. S., & Ratcliffe, J. H. (2007). A micro-spatial analysis of the demographic and criminogenic environment of drug markets in Philadelphia. *Australian & New Zealand Journal of Criminology*, 40(1), 43–63.
- McCullagh, P., & Nelder, J. A. (1993). *Generalized linear models*. London, UK: Chapman & Hall.
- McFadden, D. L. (1984). Econometric analysis of qualitative response models. In Z. Griliches, & M. Intrilligator (eds.) *Handbook of econometrics*, vol. 2, pp. 1396–1456. Amsterdam, Netherlands: Elsevier.
- McPhail, C. (1991). *The myth of the madding crowd*. New York, NY: Aldine.
- McPhail, C. (1994). The dark side of purpose: Individual and collective violence in riots. *The Sociological Quarterly*, 35(1), 1–32.
- Messner, S. F., Anselin, L., Baller, R. D., Hawkins, D. F., Deane, G., & Tolnay, S. E. (1999). The spatial patterning of county homicide rates: An application of exploratory spatial data analysis. *Journal of Quantitative Criminology*, 15(4), 423–450.
- Metropolitan Police (2011). *Operation Kirkin strategic review: Interim report*. London, UK: Metropolitan Police.
- Metropolitan Police (2012). *4 Days in August: Strategic review into the disorder of August 2011*. London, UK: Metropolitan Police.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298(5594), 824–827.

- Mohar, B. (1991). The Laplacian spectrum of graphs. In Y. Alavi, O. Chartrand, O. R. Oellermann, & A. J. Schwenk (eds.) *Graph theory, combinatorics, and applications*, pp. 871–898. New York, NY: Wiley.
- Mohar, B. (1997). Some applications of Laplace eigenvalues of graphs. In G. Hahn, & G. Sabidussi (eds.) *Graph symmetry*, no. 497 in NATO ASI Series, pp. 225–275. Amsterdam, Netherlands: Springer.
- Mohler, G. O. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493), 100–108.
- Moore, C., & Newman, M. E. J. (2000). Epidemics and percolation in small-world networks. *Physical Review E*, 61(5), 5678–5682.
- Morgan, F. (2001). Repeat burglary in a Perth suburb: Indicator of short-term or long-term risk? In G. Farrell, & K. Pease (eds.) *Repeat victimisation*, vol. 12 of *Crime Prevention Studies*, pp. 83–118. Monsey, NY: Criminal Justice Press.
- Myers, D. J. (2000). The diffusion of collective violence: Infectiousness, susceptibility, and mass media networks. *American Journal of Sociology*, 106(1), 173–208.
- Newman, M. (2003). The structure and function of complex networks. *SIAM Review*, 45(2), 167–256.
- Newman, M. (2010). *Networks: An introduction*. Oxford, UK: Oxford University Press.
- Newman, M. E. J. (2011). Complex systems: A survey. *American Journal of Physics*, 79(8), 800–810.
- Newman, M. E. J. (2013a). Spectral community detection in sparse networks. arXiv e-print 1308.6494.
- Newman, M. E. J. (2013b). Spectral methods for community detection and graph partitioning. *Physical Review E*, 88(4), 042822.
- Newman, M. E. J., & Barkema, G. T. (1999). *Monte Carlo methods in statistical physics*. Oxford, UK: Oxford University Press.

- Newman, M. E. J., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2), 026118.
- Newman, O. (1972). *Defensible space: Crime prevention through urban design*. New York, NY: Macmillan.
- Neyman, J. (1939). On a new class of "contagious" distributions, applicable in entomology and bacteriology. *The Annals of Mathematical Statistics*, 10(1), 35–57.
- North, B. V., Curtis, D., & Sham, P. C. (2002). A note on the calculation of empirical p values from Monte Carlo procedures. *American Journal of Human Genetics*, 71(2), 439–441.
- Office for National Statistics (2001). *2001 census: Aggregate data (England and Wales)*. UK Data Service Census Support.
- Office for National Statistics (2011). *Output areas (OAs)*. UK Data Service Census Support.
- Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2), 379–402.
- Okabe, A., Satoh, T., & Sugihara, K. (2009). A kernel density estimation method for networks, its computational method and a GIS-based tool. *International Journal of Geographical Information Science*, 23(1), 7–32.
- Pach, J. (2004). *Towards a theory of geometric graphs*. Providence, RI: American Mathematical Society.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web.
- Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14), 3200–3203.

- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London, UK: Sage.
- Pease, K. (1998). *Repeat victimisation: Taking stock*. London, UK: Home Office Police Research Group.
- Penrose, M. (2003). *Random geometric graphs*. Oxford; UK: Oxford University Press.
- Perry, W. L., McInnis, B., Price, C. C., Smith, S., & Hollywood, J. S. (2013). *Predictive policing*. Santa Monica, CA: RAND.
- Pitcher, A. B. (2010). Adding police to a mathematical model of burglary. *European Journal of Applied Mathematics*, 21(Special Double Issue 4-5), 401–419.
- Pitcher, A. B., & Johnson, S. D. (2011). Exploring theories of victimization using a mathematical model of burglary. *Journal of Research in Crime and Delinquency*, 48(1), 83–109.
- Porta, S., Crucitti, P., & Latora, V. (2006a). The network analysis of urban streets: A dual approach. *Physica A: Statistical Mechanics and its Applications*, 369(2), 853–866.
- Porta, S., Crucitti, P., & Latora, V. (2006b). The network analysis of urban streets: A primal approach. *Environment and Planning B: Planning and Design*, 33(5), 705–725.
- Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Cardillo, A., Belli, E., Cárdenas, F., Cormenzana, B., & Latora, L. (2012). Street centrality and the location of economic activities in Barcelona. *Urban Studies*, 49(7), 1471–1488.
- Porta, S., Latora, V., Wang, F., Strano, E., Cardillo, A., Scellato, S., Iacoviello, V., & Messori, R. (2009). Street centrality and densities of retail and services in Bologna, Italy. *Environment and Planning B: Planning and Design*, 36(3), 450–465.
- Ratcliffe, J. H. (2002). Aoristic signatures and the spatio-temporal analysis of high volume crime patterns. *Journal of Quantitative Criminology*, 18(1), 23–43.

- Ratcliffe, J. H. (2004). The hotspot matrix: A framework for the spatio-temporal targeting of crime reduction. *Police Practice and Research*, 5(1), 5–23.
- Ratcliffe, J. H., & Rengert, G. F. (2008). Near-repeat patterns in Philadelphia shootings. *Security Journal*, 21(1-2), 58–76.
- Ratti, C. (2004). Space syntax: Some inconsistencies. *Environment and Planning B: Planning and Design*, 31(4), 487 – 499.
- Rey, S. J., Mack, E. A., & Koschinsky, J. (2012). Exploratory space–time analysis of burglary patterns. *Journal of Quantitative Criminology*, 28(3), 509–531.
- Reynald, D. M. (2010). Guardians on guardianship: Factors affecting the willingness to supervise, the ability to detect potential offenders, and the willingness to intervene. *Journal of Research in Crime and Delinquency*, 47(3), 358–390.
- Reynald, D. M., & Elffers, H. (2009). The future of Newman’s defensible space theory: Linking defensible space and the routine activities of place. *European Journal of Criminology*, 6(1), 25–46.
- Riolo, M. A., & Newman, M. E. J. (2014). First-principles multiway spectral partitioning of graphs. *Journal of Complex Networks*, 2(2), 121–140.
- Riots Communities and Victims Panel (2011). 5 Days in August: An interim report on the 2011 English riots.
- Robinson, W. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3), 351–357.
- Rosvall, M., Trusina, A., Minnhagen, P., & Sneppen, K. (2005). Networks and cities: An information perspective. *Physical Review Letters*, 94(2), 028701.
- Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4), 581–603.
- Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, 94(4), 774–802.

- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277(5328), 918–924.
- Scellato, S., Cardillo, A., Latora, V., & Porta, S. (2006). The backbone of a city. *The European Physical Journal B - Condensed Matter and Complex Systems*, 50(1-2), 221–225.
- Schelling, T. C. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2), 143–186.
- Shaw, C. R., & McKay, H. D. (1969). *Juvenile delinquency and urban areas: A study of rates of delinquency in relation to differential characteristics of local communities in American cities*. Chicago, IL: University of Chicago Press.
- Sherman, L., & Eck, J. (2002). Policing for crime prevention. In L. Sherman, D. Farrington, B. Welsh, & D. L. MacKenzie (eds.) *Evidence-based crime prevention*. Abingdon, UK: Routledge.
- Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1), 27–56.
- Short, M. B., D’Orsogna, M. R., Brantingham, P. J., & Tita, G. E. (2009). Measuring and modeling repeat and near-repeat burglary effects. *Journal of Quantitative Criminology*, 25(3), 325–339.
- Short, M. B., D’Orsogna, M. R., Pasour, V. B., Tita, G. E., Brantingham, P. J., Bertozzi, A. L., & Chayes, L. B. (2008). A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences*, 18(S1), 1249–1267.
- Simpson, E. H. (1949). Measurement of diversity. *Nature*, 163, 688.
- Smith, L. M., Bertozzi, A. L., Brantingham, P. J., Tita, G. E., & Valasik, M. (2012). Adaptation of an ecological territorial model to street gang spatial patterns in Los Angeles. *Discrete and Continuous Dynamical Systems*, 32(9), 3223–3244.
- Snook, B. (2004). Individual differences in distance travelled by serial burglars. *Journal of Investigative Psychology and Offender Profiling*, 1(1), 53–66.

- Spelman, W. (1993). Abandoned buildings: Magnets for crime? *Journal of Criminal Justice*, 21(5), 481–495.
- Strano, E., Nicosia, V., Latora, V., Porta, S., & Barthélemy, M. (2012). Elementary processes governing the evolution of road networks. *Scientific Reports*, 2, 296.
- Strogatz, S. H. (1994). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Reading, MA: Perseus Books.
- Summers, L., Johnson, S. D., & Rengert, G. F. (2010). The use of maps in offender interviewing. In W. Bernasco (ed.) *Offenders on offending: Learning about crime from criminals*, pp. 246–272. Cullompton, UK: Willan.
- The Guardian and LSE (2011). *Reading the riots: Investigating England's summer of disorder*. London, UK: The Guardian.
- Tilley, N. (2009). What's the “what” in “what works?”? Health, policing and crime prevention. In J. Knutsson, & N. Tilley (eds.) *Evaluating crime reduction initiatives*, vol. 24 of *Crime Prevention Studies*, pp. 121–145. Cullompton, UK: Willan.
- Tilley, N., Pease, K., Hough, M., & Brown, R. (1999). *Burglary prevention: Early lessons from the crime reduction programme*. No. 1 in Crime Reduction Research Series. London, UK: Home Office Policing and Reducing Crime Unit.
- Tita, G., & Ridgeway, G. (2007). The impact of gang formation on local patterns of crime. *Journal of Research in Crime and Delinquency*, 44(2), 208–237.
- Tompson, L., & Bowers, K. (2013). A stab in the dark? A research note on temporal patterns of street robbery. *Journal of Research in Crime and Delinquency*, 50(4), 616–631.
- Torrens, P. M., & McDaniel, A. W. (2013). Modeling geographic behavior in riotous crowds. *Annals of the Association of American Geographers*, 103(1), 20–46.
- Townsley, M. (2008). Visualising space time patterns in crime: The hotspot plot. *Crime Patterns and Analysis*, 1(1), 61–74.

- Townsley, M., Homel, R., & Chaseling, J. (2000). Repeat burglary victimisation: Spatial and temporal patterns. *Australian & New Zealand Journal of Criminology*, 33(1), 37–63.
- Townsley, M., Homel, R., & Chaseling, J. (2003). Infectious burglaries: A test of the near repeat hypothesis. *British Journal of Criminology*, 43(3), 615–633.
- Townsley, M., & Sidebottom, A. (2010). All offenders are equal, but some are more equal than others: Variation in journeys to crime between offenders. *Criminology*, 48(3), 897–917.
- Train, K. (2003). *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press.
- Tseloni, A., & Pease, K. (2003). Repeat personal victimization: ‘Boosts’ or ‘flags’? *British Journal of Criminology*, 43(1), 196–212.
- van Mieghem, P. (2012). *Graph spectra for complex networks*. Cambridge, UK: Cambridge University Press.
- Vragović, I., Louis, E., & Díaz-Guilera, A. (2005). Efficiency of informational transfer in regular and complex networks. *Physical Review E*, 71(3), 036122.
- Wang, X., & Brown, D. E. (2012). The spatio-temporal modeling for criminal incidents. *Security Informatics*, 1(1), 1–17.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.
- Weidmann, N. B., & Salehyan, I. (2013). Violence and ethnic segregation: A computational model applied to Baghdad. *International Studies Quarterly*, 57(1), 52–64.
- Weisburd, D., Bruinsma, G. J. N., & Bernasco, W. (2009). Units of analysis in geographic criminology: Historical development, critical issues, and open questions. In D. Weisburd, W. Bernasco, & G. J. N. Bruinsma (eds.) *Putting crime in its place*, pp. 3–31. New York, NY: Springer.

- Weisburd, D., Bushway, S., Lum, C., & Yang, S.-m. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology*, 42(2), 283–322.
- Wernicke, S., & Rasche, F. (2006). FANMOD: A tool for fast network motif detection. *Bioinformatics*, 22(9), 1152–1153.
- White, G. F. (1990). Neighborhood permeability and burglary rates. *Justice Quarterly*, 7(1), 57–67.
- Wiles, P., & Costello, A. (2000). *The ‘road to nowhere’: The evidence for travelling criminals*. No. 207 in Home Office Research Studies. London, UK: Home Office Research, Development and Statistics Directorate.
- Wilson, A. G. (1970). *Entropy in urban and regional modelling*. London, UK: Pion.
- Wilson, A. G. (1971). A family of spatial interaction models, and associated developments. *Environment and Planning*, 3(1), 1–32.
- Wilson, A. G. (2008). Boltzmann, Lotka and Volterra and spatial structural evolution: An integrated methodology for some dynamical systems. *Journal of The Royal Society Interface*, 5(25), 865–871.
- Winfrey, A. T. (2001). *The geometry of biological time*. New York, NY: Springer, 2nd ed.
- Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media.
- Wortley, R. (2001). A classification of techniques for controlling situational precipitators of crime. *Security Journal*, 14(4), 63–82.
- Xu, X., Zhang, J., & Small, M. (2008). Superfamily phenomena and motifs of networks induced from time series. *Proceedings of the National Academy of Sciences*, 105(50), 19601–19605.
- Xue, Y., & Brown, D. E. (2006). Spatial analysis with preference specification of latent decision makers for criminal event prediction. *Decision Support Systems*, 41(3), 560–573.

- Ye, X., & Wu, L. (2011). Analyzing the dynamics of homicide patterns in Chicago: ESDA and spatial panel approaches. *Applied Geography*, 31(2), 800–807.
- Yiu, S. Y., Gill, A. W., & Shi, P. (2003). Using agent based distillations to model civil violence management. *Journal of Battlefield Technology*, 6(1), 27–32.